Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 153574
UNCLASSIFIED	

TITLE

DEFINITION OF A PERFORMANCE EVALUATION METHODOLOGY FOR SENSOR DATA FUSION SYSTEMS

System Number: 153574

Patron Number: 8242

Requester: . DTIC

Defense Technical Information Cent
8725 John J. Kingman Road, Sui

Fort Belvoir VA
22060-6218
United States

Notes:

DSIS Use only:

Deliver to: TC

Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 153574
UNCLASSIFIED	

TITLE

DEFINITION OF A PERFORMANCE EVALUATION METHODOLOGY FOR SENSOR DATA FUSION SYSTEMS

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to:

Défense nationale National Defence

UNCLASSIFIED

DEFENCE RESEARCH ESTABLISHMENT
CENTRE DE RECHERCHES POUR LA DÉFENSE
VALCARTIER, QUÉBEC

RESEARCH AND DEVELOPMENT BRANCH
DEPARTMENT OF NATIONAL DEFENCE
CANADA
BUREAU - RECHERCHE ET DÉVELOPPEMENT
MINISTÈRE DE LA DÉFENSE NATIONALE

Canada

SANS CLASSIFICATION

UNCLASSIFIED

DEFENCE RESEARCH ESTABLISHMENT
CENTRE DE RECHERCHES POUR LA DÉFENSE
VALCARTIER, QUÉBEC

DREV - R - 9423

Unlimited Distribution / Distribution illimitée

DEFINITION OF A PERFORMANCE EVALUATION METHODOLOGY FOR SENSOR DATA FUSION SYSTEMS

by

J. Roy and É. Bossé

September / septembre 1995

Approved by / approuvé par

Director / Directeur

14/8/95
Date

SANS CLASSIFICATION

ABSTRACT

A key element in the anticipated information management problem on a naval platform is the ability to combine or fuse data, not only as a volume-reducing strategy, but also as a means to exploit the unique combinations of data that may be available. In this regard, the Command and Control Information Systems Division at the Defence Research Establishment Valcartier (DREV) is involved in multiple R&D activities in the field of local area Multi-Sensor Data Fusion (MSDF) for naval command and control afloat. The ultimate goal of these activities is to develop an optimal MSDF function that will enhance the tactical performance of the command and control system of the Canadian Patrol Frigate. In this context, rigorously evaluating the performance of MSDF systems is of prime importance. The system designer may have several MSDF concepts among which to choose in order to fulfill a particular requirement, the MSDF system may require optimization to improve performance, or the MSDF system may be undergoing testing to assure that it is operating correctly. Unfortunately, no widely accepted scheme for characterizing the performance of MSDF systems is currently in use. This document proposes such a scheme (or methodology) to be used for MSDF performance evaluation in computer simulations.

RÉSUMÉ

Un élément fondamental du problème anticipé de la gestion d'information sur une plate-forme navale est la capacité de combiner ou fusionner les données, non seulement comme stratégie de réduction du volume d'information, mais aussi comme moyen d'exploiter les combinaisons uniques de données qui peuvent être disponibles. À cet égard, la Division des systèmes d'information du commandement et contrôle au Centre de recherches pour la défense, Valcartier (CRDV) participe à plusieurs activités de R&D dans le domaine de la fusion locale de données multi-capteurs pour le commandement et contrôle naval au large. Le but ultime de ces activités est de développer une fonction optimale de fusion de données qui améliorera la performance tactique du système de commandement et contrôle pour la Frégate de patrouille canadienne. Dans ce contexte, il est primordial de pouvoir évaluer rigoureusement la performance des systèmes de fusion de données. Le concepteur de systèmes peut avoir à choisir parmi plusieurs concepts de fusion de données dans le but de combler un besoin particulier. Le système de fusion de données peut nécessiter une optimisation pour améliorer sa performance, ou bien on peut faire des tests pour s'assurer qu'il fonctionne correctement. Malheureusement, il n'existe aucune approche généralement reconnue pour caractériser la performance des systèmes de fusion de données. Ce document propose une telle approche (ou méthodologie) pour une utilisation dans les évaluations de la performance des systèmes de fusion de données pour les simulations sur ordinateur.

PRECEDING PAGE BLANK

TABLE OF CONTENTS

TABLE OF CONTENTS (contd)

EXECUTIVE SUMMARY

Current Above Water Warfare (AWW) systems are largely composed of stand-alone subsystems. Until recently, there has been a tendency for various AWW sensors and weapons to be conceived, developed and produced in isolation from each other. These components are only to be brought together on the ship and there, superficially integrated. However, the development and/or acquisition of advanced AWW sensor and weapon elements, although necessary, are not sufficient for providing the required protection of ships against the anticipated future threats. The simple interfacing of these elements is not enough because such independent AWW components are seldom used in a coordinated manner. This typically leads to a confusing and time-late decision environment for the ship's commander. Hence, the effectiveness of the AWW system is not only determined by the capabilities of the AWW sensor and weapon suites alone, but also by the effectiveness of the AWW system integration which must focus on cooperative, synergistic and efficient utilization of all of the AWW sensor and weapon elements.

In this regard, the Command and Control Information Systems Division at Defence Research Establishment Valcartier (DREV) is involved in multiple R&D activities in the field of local area Multi-Sensor Data Fusion (MSDF) for naval command and control afloat. The ultimate goal of these activities is to develop an optimal MSDF function that will enhance the tactical performance of the command and control system (CCS) of the Canadian Patrol Frigate (CPF). In this context, rigorously evaluating the performance of MSDF systems is of prime importance. The system designer may have several MSDF concepts among which to choose in order to fulfill a particular requirement; the MSDF system may require optimization to improve performance; or the MSDF system may be undergoing testing to assure that it is operating correctly. Unfortunately, no widely accepted scheme for characterizing the performance of MSDF systems is currently in use. This document proposes such a scheme (or methodology) to be used for MSDF performance evaluation in computer simulations.

The MSDF performance evaluation methodology described in this document provides an appropriate framework to guide the MSDF system designers in the evaluation of current and future integrated surveillance and tracking systems suitable to fulfill the Canadian Forces requirements, and in the optimization of the operation of these systems to obtain the best performance.

LIST OF ACRONYMS

AAW	Anti-Air Warfare
AWW	Above Water Warfare
C^2	Command and Control
CASE_ATTI	Concept Analysis and Simulation Environment for Automatic Target Tracking and Identification
CCIS	C^2 Information System
CCS	C^2 System
CPF	Canadian Patrol Frigate
CRAD	Chief, Research and Development
CRDV	Centre de recherches pour la défense, Valcartier
DND	Department of National Defence
DREV	Defence Research Establishment Valcartier
EMCON	Emissions Control
E-O	Electro-Optical
ESA	Electronically Scanned Antenna
ESM	Electronic Support Measure
GPS	Global Positioning System
INS	Inertial Navigation System
I/O	Input/Output
MHT	Multiple Hypothesis Tracking
MOE	Measure Of Effectiveness
MOFE	Measure Of Force Effectiveness
MOM	Measure Of Merit
MOP	Measure Of Performance
MSA	Mechanically Scanned Antenna
MSDF	Multi-Sensor Data Fusion
R&D	Research and Development
SS/MTTI	Single Sensor / Multiple Target Tracking and Identification
TWS	Track-While-Scan
VOI	Volume Of Interest

1.0 INTRODUCTION

A research project has been undertaken by the CCIS Division at the Defence Research Establishment Valcartier (DREV) to investigate in depth issues related to Multi-Sensor Data Fusion (MSDF). The objective of this project is to analyze, evaluate and develop advanced techniques to automatically produce the best possible estimate of the position, kinematic behavior, and identification of all objects surrounding a single ship, mainly through the fusion of data from dissimilar organic sensors (e.g., radar, E-O, ESM), while including inorganic information (e.g., data coming over communication links, intelligence reports, etc.). The use of the latter type of information is directed towards the potential enhancement of the performance of the different sensor data fusion subprocesses. The end result of MSDF (i.e., a highly reliable computation of the tactical picture) is used as an input to the subsequent, higher level situation assessment and threat evaluation C^2 processes.

An overview of the R&D activities involving the Data Fusion group at DREV in the field of local area MSDF for naval command and control afloat, and a description of the CASE_ATTI (Concept Analysis and Simulation Environment for Automatic Target Tracking and Identification) simulation testbed, that has been developed to support the theoretical work, were already presented in previous documents (Refs. 1-2). As can be drawn from these documents, the ultimate goal of the MSDF project is to develop an optimal MSDF function that will enhance the tactical performance of the command and control system (CCS) of the Canadian Patrol Frigate (CPF).

In this regard, the CASE_ATTI system is currently being used to support the development of advanced MSDF concepts that could apply to the current CPF sensor suite, as well as its anticipated upgrades, in order to improve its AWW performance against predicted future threats. This very practical study aims at identifying and developing techniques for combining radar/EO/ESM data, and at evaluating the real benefits of the combination. Two major aspects need to be addressed for this application: first, the representation of the actual CPF sensor suite to establish its baseline performance and second, the quantification of the performance improvements gained when using an upgraded sensor suite combined with advanced MSDF concepts.

In this context, rigorously evaluating the performance of MSDF systems is of prime importance. In particular, the system designer may have several MSDF concepts among which to choose in order to fulfill a particular requirement; the MSDF system may require tuning for improved performance; or the MSDF system may be undergoing testing to assure that it is operating correctly (Ref. 3). If the appropriate data on the performance of weapon systems are available, various analyses can be conducted about trade-offs between improvements in information performance (due to MSDF) and weapon performance (Ref. 4). For instance, should one invest in MSDF to extend a combat direction system's detection range, or should those same dollars be spent to increase the weapon's kill probability through warhead improvement?

Unfortunately, no widely accepted scheme for characterizing the performance of MSDF systems is currently in use (Ref. 5). While much research is being performed to develop and apply new MSDF algorithms and techniques, little work has been performed to determine how well such methods work or to compare alternative methods against a common problem. This document defines a methodology to be used for MSDF performance evaluation in computer simulations.

The document is organized as follows. Chapter 2 discusses the use of sensors and advanced data fusion systems in naval defence applications in the perspective of evaluating their performance for surveillance and tracking. The MSDF system is then discussed (in terms of its internal subprocesses, environment and output results) in order to establish a solid basis for the rest of the document.

Chapter 3 addresses some of the performance evaluation issues encountered in the use of MSDF systems in defence applications. In particular, this chapter discusses some of the reasons why MSDF systems are hard to evaluate. It also discusses the prediction of tracking performance using analytical evaluation methods. It very briefly addresses covariance analysis for use in design and evaluation of the tracking filters, expressions for use in the preliminary determination of observation-to-track correlation performance, and finally, Markov chain techniques for use in determining target track maintenance and false track statistics. This chapter ends with discussions on the importance of Monte Carlo simulation techniques, and on experimental trials for performance evaluation.

A framework for the definition and discussion of the MSDF performance evaluation process is suggested in Chapter 4. This proposed methodology for evaluating and comparing alternative MSDF systems in a variety of common target environments comprises seven essential steps which are described in this chapter.

The research and development activities described in this document were performed at DREV between July and October 1994 under PSC 12C, Ship Combat System Integration.

2.0 MULTI-SENSOR DATA FUSION SYSTEMS

This chapter discusses the use of sensors and advanced data fusion systems in naval defence applications in the perspective of evaluating their performance for surveillance and tracking. These applications generally involve detecting the presence of an unknown number of objects of interest (generally referred to as "targets"), some of which may be hostile, some friendly and some neutral, and estimating (or tracking) their position, motion and identity from periodic sensor measurements.

Figure 1 illustrates the main components of the MSDF performance evaluation concept. The ground truth picture represents the real composition and status of a scenario of tactical interest. The objective of the MSDF system, whose performance is affected by factors such as energy propagation effects, is to keep an awareness of the external situation as efficiently and accurately as possible (Ref. 6). During its operation, the MSDF system generates both a measured tactical picture (i.e., the output, not explicitly shown on Fig. 1, of the sensing process) and an estimated picture (whose quality should in principle be superior to that of the measured picture). The purpose of the performance evaluation process is to evaluate the ability of the MSDF system to generate measured and estimated tactical pictures that accurately reproduce the ground truth tactical picture. The evaluation objective is to quantify the potential divergence that may result from the many limitation factors affecting the performance of MSDF systems.

In this chapter, we discuss the MSDF system (in terms of both its internal subprocesses and output results) and its environment in order to establish a solid basis for the rest of the document.

2.1 Ground Truth Tactical Picture

The ground truth tactical picture serves as a basis for comparison with the measured and estimated tactical pictures. It depicts the known activities (i.e., position, kinematic behavior, emissions, and identity) of a known number N_g of real, distinct targets in a given area of interest. A target may be a plane, ship, missile, etc. Targets possess defined kinematic and non-kinematic properties. The trajectory of a target typically summarizes its kinematic properties. The target type or category, its allegiance, nationality, threat level and specific identification are examples of non-kinematic

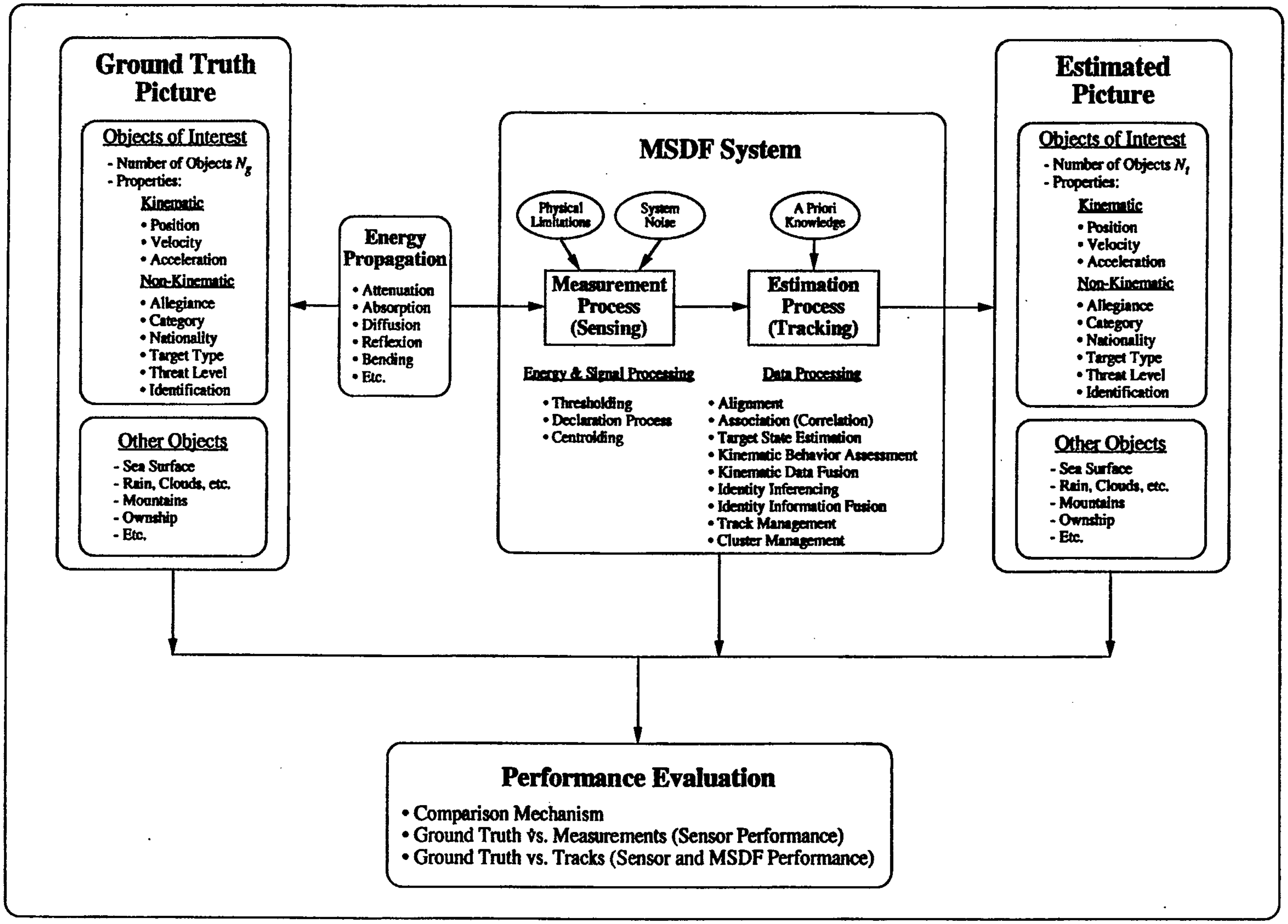

FIGURE 1 - Performance evaluation concept for MSDF systems

properties. The ground truth targets progress in time and space according to a well-defined scenario that also includes the characteristics of the target emissions.

The MSDF system has to operate in the four-dimensional world of space and time. Typically, the commander of a ship must assess all the ongoing activities within a given space volume surrounding the ship. This volume is referred to as the Volume Of Interest (VOI). For an MSDF system, it is defined by mission considerations, and there is no guarantee that the VOI will correspond to the MSDF system's sensor coverages (Ref. 7). For example, the mission might be fleet defence against cruise missiles, and for various reasons the MSDF system may not be able to "see" or recognize all of the launch platforms within cruise missile range. The VOI for the MSDF system might be static. However, it might also be mobile and fluid, e.g., an elongated bubble centered on a distant threat, or a defensive perimeter enclosing friendly forces. If the VOI is mobile, it carves out a swath, and the MSDF system can be expected to "remember" targets that have departed it or been left behind the VOI.

In the development of performance evaluation plans for an MSDF system, one must determine which targets have penetrated its VOI, so that one can appraise its performance in finding them. Test scenarios for an MSDF system must be larger than the VOI, so that targets will, over the duration of the scenario, both enter and exit the VOI.

The VOI for the estimation process only is defined by those, and only those targets actually detected by the sensors. For the estimation process, our assumptions about sensor coverage and activity, coupled with the scenario, implicitly define the VOI. One must, in the estimation process performance evaluation plan, determine which targets are detected. Hence, test scenarios for the estimation process only may, but do not have to be larger than the VOI.

Figure 2 is an illustration of the concept of VOI discussed above. Since one is almost certainly concerned with a dynamic environment, one needs to take time into account as well. This is also shown in Fig. 2. The time interval of interest for performance evaluation is any suitable interval during which one is interested in comparing the measured and estimated tactical pictures with the ground truth picture. Typically, snapshots of these pictures at some given time "k" of interest will be used for

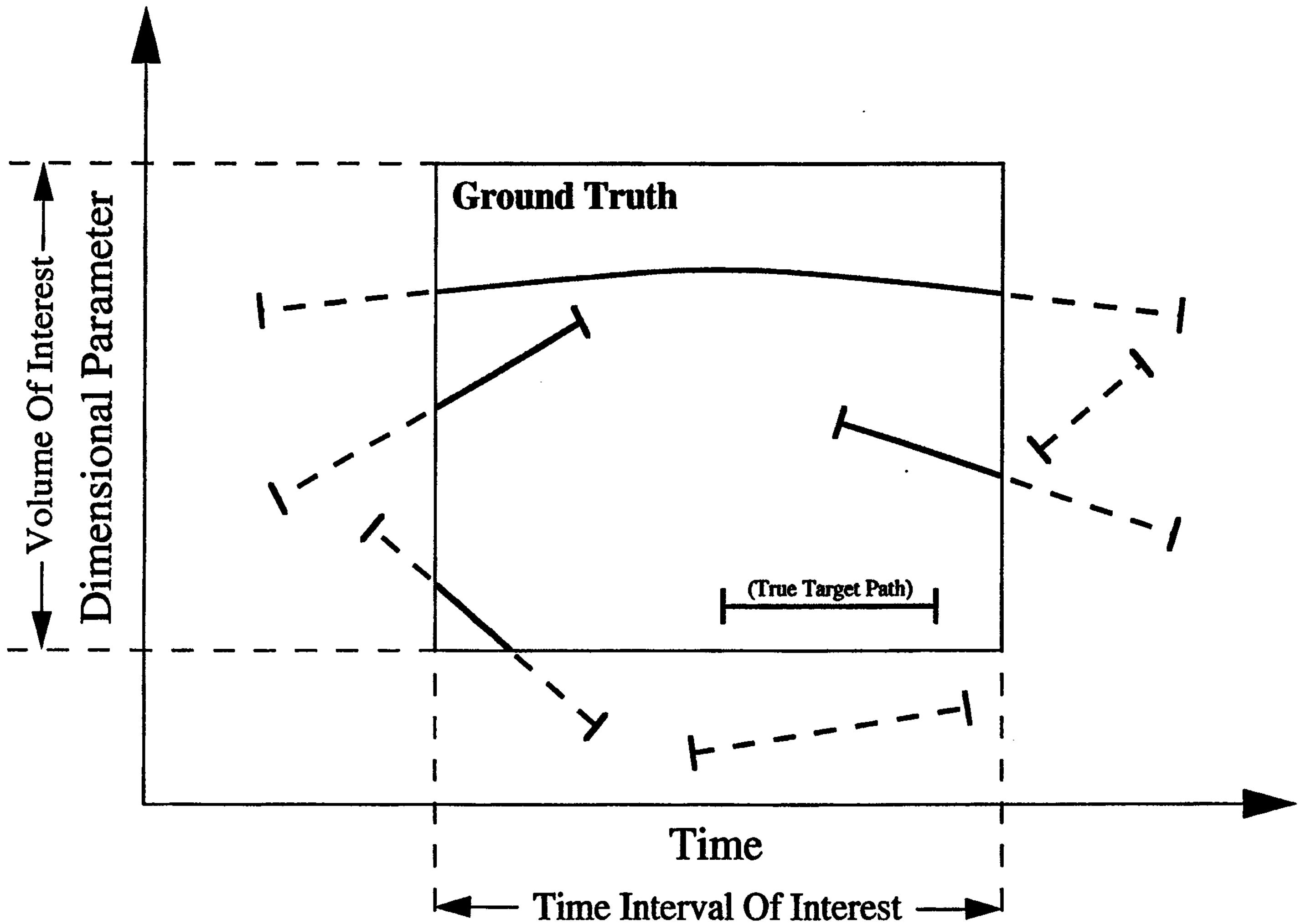

FIGURE 2 - Illustration of the concept of volume and time interval of interest

the comparison. The overall time interval of interest will thus be covered using a number of such snapshots.

Any realistic physical tracking environment also includes a number of unwanted objects of no immediate interest, e.g., sea surface, ground, mountains, birds, insects, clouds, rain and other meteorological phenomena, etc. These objects may cause returns that "clutter" the sensor display with false target declarations that may overload the processor elements and/or desensitize the sensor to the true targets. Although the important performance degradation effects potentially introduced by these unwanted objects must be taken into account by the MSDF designer, they will not be given further considerations with respect to performance evaluation in the rest of this document.

Hence, for the purpose of MSDF system performance evaluation in computer simulations, reality or ground truth is represented by only the N_g objects of interest progressing within the VOI during the time interval of interest. Using a formal notation similar to the one used in Ref. 7, it is assumed that the MSDF evaluator knows at any time k all the characteristics and behavior of the set of targets

$$G(k) = \left\{ g_1, \dots, g_{N_g(k)} \right\} \qquad [1]$$

which are actually part of ground truth.

2.2 Sensors and MSDF system

The objective of an ideally effective MSDF system is to establish a number of clean, stable tracks that corresponds exactly to the number of objects in the physical environment. As illustrated in Fig. 1, it is assumed in this document that the overall tracking system (or MSDF system) comprises a measurement (or sensing) component that provides observations of the target environment, and an estimation (or tracking) component (ranging from a simple software tracking filter to a sophisticated multi-target multi-sensor data fusion system), which

1) acquires and maintains unambiguous, stable tracks corresponding to the perceived population of real objects within the volume of interest (Ref. 5),

2) estimates the state and identity of each tracked object with the objective of keeping an accurate and complete awareness of the external environment (Ref. 6), and
3) suppresses clutter and other unwanted objects (i.e., discards "uninteresting" targets from the scene).

There is no particular requirement that the measurement and tracking components be collocated. Figure 3 illustrates the usual definition of three types of MSDF architecture for two generic sensors (Refs. 1 and 4). One possible type of MSDF architecture is based on maintaining sensor-level tracks at each sensor site, finding the sensor tracks that potentially represent the same target and then combining these tracks into global tracks of the MSDF function. A second type of architecture assumes that the raw sensor measurements (i.e., sensor contacts) are sent directly to the MSDF function to be combined into global tracks. This architecture is sometimes referred to as a "central-level" architecture since the tracks are only formed in the central processor. Finally, fusion at the signal-level typically combines signals from similar sensors to produce a better quality signal of the same form.

The MSDF architecture is an important issue since the expected benefits are different depending on the way the sensor data is combined. The selection of the MSDF architecture type should be aimed at optimizing the target detection, tracking and identification performance required for a specific ship given its missions. However, the selection is also constrained by the technological capabilities (hardware and software) of the sensors and the C^2 system (CCS).

2.2.1 Measurement Process

The measurement process encompasses both the energy and signal processing aspects. As part of signal processing, target detection is the process of determining the presence of a target, usually by declaring a target present if a voltage exceeds a threshold. Typically, military sensors generate detections that involve "point" targets (i.e., threshold crossings in only a few resolution cells), from which certain targets properties are measured (Ref. 6). This is discussed in more depth below.

A sensor will typically spend a limited amount of time on a single target because, in most cases, scanning is necessary in order to provide updated information on

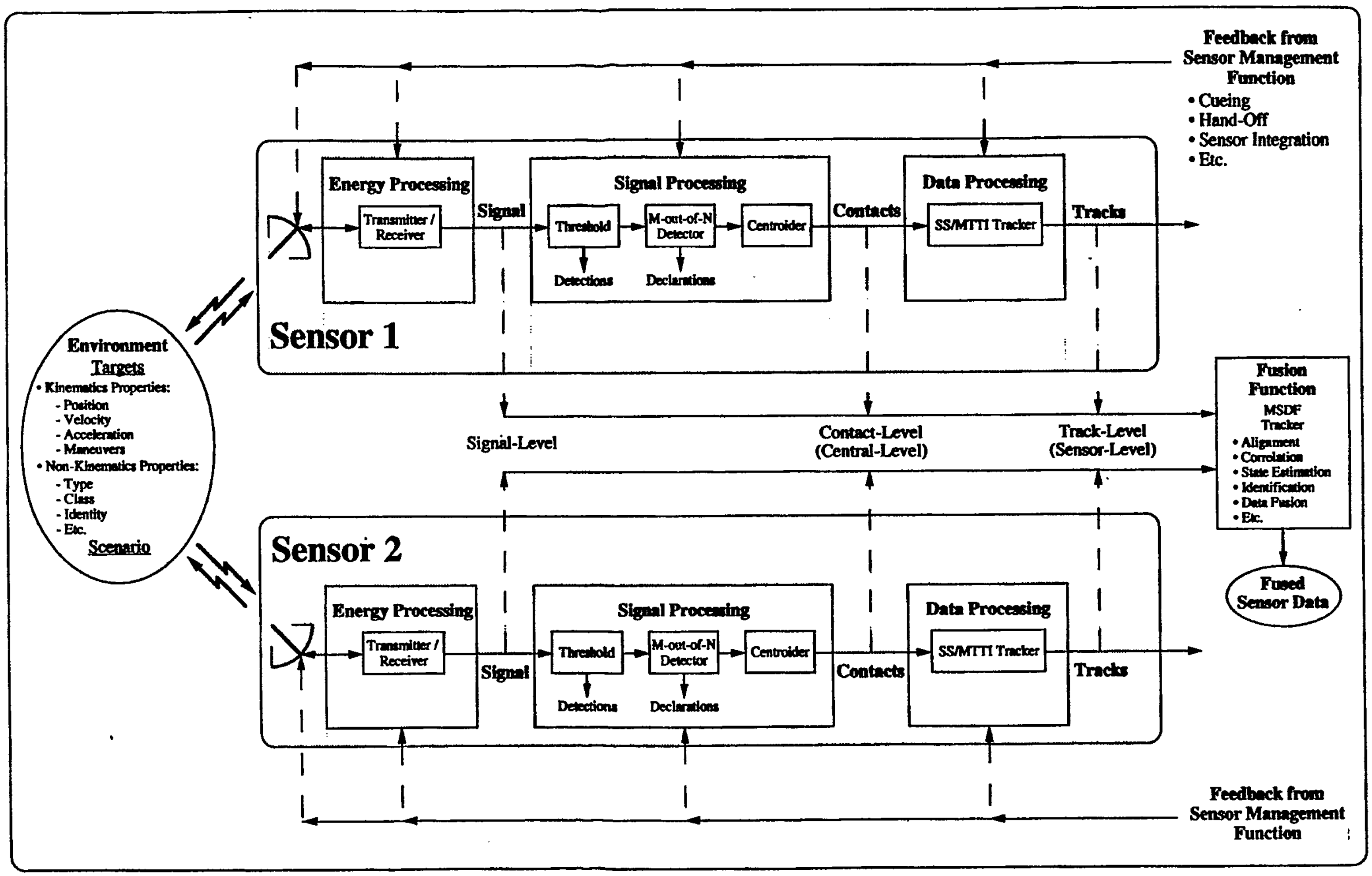

FIGURE 3 - Definition of three types of MSDF architecture for two generic sensors

established tracks and to search for new targets. One important sensor design consideration is the selection of a decision rule on the return received during the time on target, so as to discriminate between returns from targets of interest and returns from extraneous sources such as clutter. A widely spread approach to this decision process is to compare the incoming signal power to a threshold which is typically set so that the probability of false alarm (P_{FA}) remains constant. A "detection" occurs each time the received power exceeds the selected threshold.

For a given threshold setting, the probability of target detection (P_D) will generally be a complex function of the sensor capabilities, the target size, the sensor-target geometry, and the physical environment (atmospheric attenuation, etc.). The threshold value (and resulting P_D and P_{FA}) should be selected taking into account its effect on overall tracking system performance. It may even be desirable to set the threshold adaptively.

Typically, it is assumed that the measurement set produced by a given sensor during a single scan contains at most one observation from each target which may be within the search volume of this sensor. This may require some redundancy elimination logic in the measurement preprocessing so that multiple simultaneous detections from the same source are combined. Typically, a sensor search volume may be covered using two or more bars in which the sensor scans in azimuth angle while maintaining a fixed elevation angle for each bar. In such a case, a redundancy elimination logic is required to ensure that detections received from the same target on multiple bars are not interpreted as being the result of multiple targets. The result of the centroiding of simultaneous detections to form a single report of a target is often called a "contact" (Fig. 3).

In addition to combining multiple detections from a single target, it is also desirable to recognize when a single observation was produced by multiple targets. For example, radar measurement techniques might not be able to resolve several closely spaced targets that are within the radar's beamwidth. However, radar data processing techniques have been developed to determine when there are multiple targets within the radar's beamwidth, even if distinct measurements from all targets cannot be obtained.

2.2.1.1 Advanced Measurement Process Concepts

In the traditional elaboration, MSDF is portrayed as a purely passive, open loop process (i.e., a function that simply processes whatever it receives). In the more advanced and enlarged sense, however, an MSDF system also includes many additional functions, the most essential of which is active feedback. An MSDF system not only detects, localizes, and identifies targets, but also, on the basis of an evolving picture, manages the information it might receive by pointing, focusing, maneuvering, and adaptively selecting the modalities of its sensors and sensor platforms (Ref. 7).

In order to maximize target information and fire support obtained by sensors while minimizing the threat to system and assets (e.g., through an emission control (EMCON) policy), sensor management is a decision process which addresses the following questions:

- When to search, track, remain covert?
- What regions to search, what objects to track?
- How long to search, track, remain covert?
- With which sensor combinations? Fire support?

Sensor management is thus a resource allocation problem. Sensor cueing, hand-off and scheduling issues are part of the sensor management domain.

The management of sensors may require that different sensors cooperate to acquire measurements on a common target. The two primary cooperative functions are cueing and hand-off. *Cueing* is the process of using the detections (i.e., contact-level cueing) or tracks (i.e., track-level cueing) from one sensor (A) to point another sensor (B) toward the same target or event. *Hand-off* occurs when sensor A has cued sensor B for transferring surveillance or fire control responsibility from A to B.

Two processes must occur for cueing or hand-off: (1) the cueing sensor must provide the cued sensor data that contains sufficient information to point to the target and identify it as the specific target being cued, (2) the cued sensor must search for the target of interest and verify that it has been acquired.

There are typically an ownship inner "sensor management" feedback loop, and an ownship outer "platform management" feedback loop. Multiple ownships may be operated as a next higher level unit, called a Battle Group. The steering commands from next higher level are mixed with those at the local level, to achieve both local and global objectives. Actions of friendly forces feed back into the tactical situation, producing adversarial response and new situations (Ref. 7).

Another issue related to active feedback is sensor integration that involves the modification of the sensor design so that it can receive and use pertinent information from other sensors or from the command and control process, to improve or refine its own performance. In other words, sensor integration allows the sensor to do its task better than as a stand alone autonomous sensor. Sensor integration is to a large extent a sensor system designer's issue.

In summary, an advanced MSDF system indicates what the targets are, where they are, where they aren't, and where it hasn't looked (Ref. 7). In this regard, sensors and sensor platforms are selectively employed to:

- look for, and find targets within a specified volume of interest (this implies that one demarcates the VOI, and defines search stratagems for achieving the best possible sensor coverage),
- enhance accuracy (by point and dwell) against priority targets,
- increase detections (by more frequent visits) in interesting or threatening regions,
- balance these objectives in accordance with the mission declaration, and
- operate its sensors within power, time, mutual interference, and EMCON constraints.

In practice, most current-day MSDF systems achieve feedback through the agency of operators equipped with tactical decision aids. These functions can be called "Value/Cost Analysis, Decision and Command". With the introduction of next-generation sensors, characterized by receiver-transmitter agility, abundant modalities, and multifarious constraints, automation of the sensor management and integration loop will become a virtual necessity (Ref. 7).

2.2.2 Estimation Process

As illustrated in Figs. 1 and 3, the estimation component of an MSDF system is a multilevel, multifaceted process dealing with the alignment, association or correlation, estimation, inferencing and combination of data and information from multiple sensors to achieve refined state and identity estimation for each target in the tracking environment.

Strictly speaking, estimation is the process of inferring, in some optimal fashion, the value of a parameter of interest from indirect and inaccurate observations related in a specified way to this parameter. An optimal estimator is indeed a computational algorithm that processes measurements to obtain a "best estimate" (minimum error in some sense) of a given variable of interest. This variable of interest can be:

- a parameter, a time-invariant quantity (a scalar or a vector),
- the state of a dynamic system (usually a vector).

The achievement of an estimate of the state of a system utilizes a priori information (or static inputs) such as:

- knowledge of system and measurement dynamics,
- assumed statistics of system noises and measurement errors,
- initial conditions,

so that optimal estimators are sensitive to erroneous a priori models and statistics.

The heart of any tracking system, i.e., target state estimation, is the processing of noisy data obtained from a target (typically measurements) in order to maintain an estimate of its current state (i.e., a track), which typically consists of:

1) kinematic components (position, course and speed, acceleration, etc.), and
2) non-kinematic components (target type, class, identity, radar cross-section, IR signature, etc.).

In principle, the estimates should be a more accurate assessment of the target's properties then the raw measurements. State estimation may consist in filtering (estimating the properties at the time of the latest observation), smoothing (estimating the properties at a point in the past), and prediction (estimating the properties at a point in the future).

In any MSDF system, sensor data alignment in time and space must take place before state estimation can be performed. Moreover, in order to estimate and remove the effects of sensor motion from the received data, various Inertial Navigation Systems (INS) are used, involving a wide variety of motion sensors including gyroscopes, accelerometers, and the Global Positioning System (GPS) (Ref. 6). The motion corrected observations are processed to form tracks.

The functions of data association (labeling measurements from different origins and/or sensors, at different times, that correspond to the same object or feature) and data fusion (combining measurements from different times and/or different sensors) are also required in one form or another in essentially all multiple sensor fusion applications: one function determines what information should be fused, the other function performs the fusion (Ref. 6).

The target identification aspect also needs to be considered in order to produce the complete tactical picture required by the subsequent C^2 processes. The estimation process must accurately integrate the distinguishing attributes of the targets actually observed, and provide estimates of their identification.

Finally, the tracking system must accurately indicate the correct number of targets present, and failing this, a user specified bias toward either spurious or overlooked targets. Indicating the correct number implies an ability to recover from interrupted tracks and false alarms, which in turn implies a priori knowledge of target density.

In the typical scenarios we are interested in, there can be anywhere from a few to hundreds of targets to follow. For each target, the system attempts to maintain its location, its velocity, and several types of attribute and identity estimates. For each estimate, one is interested in the error envelope, and the degree of certitude about each track. It is important that the tracker strikes a balance between too many (false) and too few (faint but real) targets. If due to the correlation or inferencing ambiguities significantly different interpretations of the data are possible, one wants to identify the differences and understand their relative likelihood (Ref. 7).

2.3 Performance Limitation Factors

Table I lists some of the many factors that make the measurement and estimation of the tactical picture by MSDF systems difficult.

The observations generated by the sensing process and passed to the tracker are affected by the characteristics of: the targets of interest, other objects in the field of view, background clutter, environmental phenomena between sensor and objects/background, sensor location relative to the objects and background, design of the sensors and signal processor algorithms, etc. The measurements may also be distorted by sensor pointing and location (navigation) errors (i.e., there are errors, called misalignments, in knowledge of the relative position and attitude of different sensors, particularly if sensors are moving independently (different platforms) (Ref. 6)). The resulting measurement vectors thus reflect feature inferencing errors, kinematic measurement errors that typically are not Gaussian, false and missing observations, impact of background clutter, unresolved closely spaced objects (also called clumps), etc.

A number of uncertainties also affects the estimation process. Target state estimation algorithms typically use some practical models of target motion in order to estimate the present and future target kinematic quantities. These target kinematic models are generally simple (such as straight-line paths, circles, etc.) and assumed to be described by well-known physical laws (e.g., ballistic laws, etc.). Unexpected changes to these assumed target motion models (i.e., "random" acceleration) are called maneuvers. Any mismatch, during a maneuver, between the real kinematic behavior of a target and the motion model assumed by a simple target state estimation algorithm can completely degrade the performance of the estimation technique. For example, in addition to the development of a track bias, some measurements actually from an object of interest can be misclassified as being from a different object, or as being noise or clutter.

The limitation factors discussed in this section are at the origin of the potential divergence between the measured and estimated tactical pictures discussed below and the ground truth tactical picture.

TABLE I

Examples of factors affecting the performance of sensors and MSDF systems

1	**Target Characteristics**	• Target Signature (e.g., Radar Cross Section) • Target Scintillation Characteristics (Variable) • Target Motion (Maneuvers)
2	**Scenarios**	• Sensor-Target Geometry • Scenario Density (Large Number of Objects) • Highly Dynamic Environment • Object Proximity (Closely Spaced Objects) - Parallel Paths - Crossing Paths
3	**Physical Environment**	• Propagation Effects (Complementary Sensors) • Clutter (Non-Zero-Mean Interference) • Enemy Interference and Spoofing
4	**Measurement Process**	• Sensor Power • Sensor Sensitivity • Sensor Detection Threshold • Sensor Measurement Resolution Capabilities • Sensor Measurement Accuracy • Sensor Noise (False Alarms) • Sensor Misalignment • Unplanned Sensor Failures • Limited Processing Capabilities
5	**Estimation Process**	• Bad Assumptions in Target Models • Bad Tuning of the Many Parameters • Limited Processing Capabilities

2.4 Measured Tactical Picture

As illustrated in Fig. 4, the snapshot output of a suite of sensors (or more precisely the output of a sensing process) to be scored at time k is assumed to be a list $M(k)$ of $N_m(k)$ measurements

$$M(k) = \{m_1, m_2, \ldots, m_{N_m(k)}\} \qquad [2]$$

that constitutes the measured tactical picture M at time k. The term measurement usually refers to a physical observation of a parameter (i.e., a parameter plus noise). In the sensor data fusion domain, measurements are thus noise-corrupted observations related in a specified way to the state of a target. Measurement is a collective term that is used to refer to all the observed (or measured) quantities included in a raw report (or a contact) output from a sensor. A measurement differs from an estimate (or a track) because an estimate operates on multiple measurements over time to extract a more accurate assessment of the parameter.

Sensor measurement characteristics must be defined. This involves specifying measurable parameters, the accuracy associated with each measured parameter, and an update interval (or an adaptive update policy) for each sensor.

In general, an observation may contain measured kinematic properties, such as position or Doppler (range rate), and measured non-kinematic properties (or attributes) such as target emitter type, radar cross section, allegiance, etc. An observation should also contain an estimate of the time at which the measurement was obtained (i.e., a time tag).

Measurement accuracies are typically specified as error variances or covariances. Measurement models generally involve non-linear functions of true object kinematics and features in assumed additive, zero-mean and temporally white noise, independent from sensor to sensor (Ref. 6).

In general, observations may be received at regular intervals of time (scans or data frames), or they may occur irregularly in time. However, because the most common occurrence is for reception at regular time intervals, we will primarily refer to observations received on scan "k" (as in Fig. 4) or with sampling interval "T". The radar

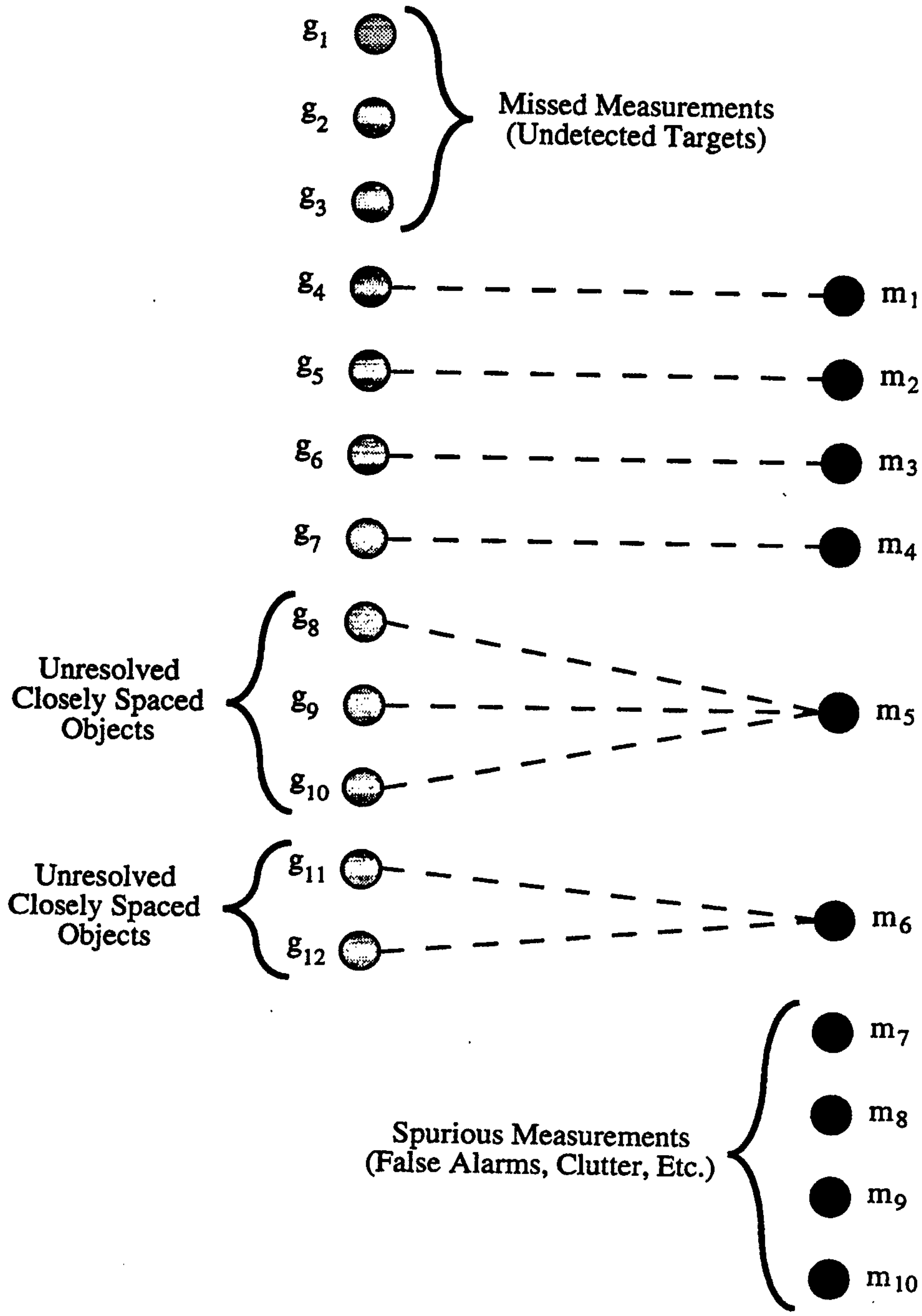

FIGURE 4 - Relationships between the ground truth (g) and measured (m) pictures (frame "*k*")

track-while-scan (TWS) system is an example of a system in which data are received at regular intervals as the radar scans a predetermined search volume and all observations are given as output at the end of each scan. Older radar systems that employed a mechanically scanned antenna (MSA) were effectively constrained to use TWS for target tracking. The more recently developed electronically scanned antenna (ESA) radars can conveniently switch back and forth between the functions of searching for new targets and illuminating existing target tracks. However, blocks of data can still be indexed by a number "k", but with the provision that the time interval between "scan k" and "scan $k+1$" may not be the same for all "k" (Ref. 8).

A "clean" measurement is of the highest quality; it corresponds to a single object of interest in the environment. Typically however, as a result of the many perturbing factors discussed in the previous section, the measured tactical picture is not composed of solely clean measurements. As illustrated in Fig. 4, there are generally divergences between the measured picture and the ground truth tactical picture. A measurement classification (taxonomy) is defined below and used to identify the various types of mismatches.

A sensor can "observe" a region containing an object but fail to detect it (i.e., object detection is not guaranteed ($P_D < 1$)), whether or not it is being tracked. Such undetected targets correspond to "missed" measurements.

Sensor measurement techniques might not be able to determine when there are several closely spaced targets within the sensor's beamwidth. In such a situation, distinct measurements from all targets cannot be obtained, and a single observation is typically produced by the multiple targets. The measurement produced by unresolved closely spaced objects is often called a "merged" measurement or a "clump".

Finally, some spurious measurements (or spurious threshold crossings) are due to noise alone (i.e., false alarms), or due to non-zero-mean interference with unknown spatial and temporal covariance (i.e., clutter). These spurious contacts can be misclassified later by the estimation process as being from an object of interest.

2.5 Estimated Tactical Picture

As illustrated in Fig. 5, the perception of truth by the MSDF system is embodied in the N_t tracks that are established or continued as the sensors sample the environment over the time interval of interest. A track t is assumed to be a triple comprising:

1) One or more state vectors estimating the target kinematic properties (i.e., position, velocity, acceleration, etc.) in a coordinate system common to all sensors, with a covariance matrix for each state vector. If more than one state vector is used, the relative likelihood (or weight) of each one is also included.
2) One or more propositions about target non-kinematic properties (attributes or identity), each with its associated likelihood function.
3) The probability of the track. It is the MSDF algorithm's estimate of the absolute likelihood that the track t exists (i.e., actually corresponds to some ground truth target).

A track is established, and the target properties are estimated, from a set of measurements (i.e., contacts) hypothesized as arising from the same object or target. The information kept in a track file also includes time tags for valid time and last update time, along with various flags which signify target maneuvers, sensor blip/scan, etc.

In the most interesting case where the Multiple Hypothesis Tracking (MHT) algorithm is used (Ref. 9), the output at time k of an advanced MSDF system to be scored is assumed to be a list

$$O(k)=\left\{\left(h_1, p_1\right), \left(h_2, p_2\right), \ldots, \left(h_{N_h(k)}, p_{N_h(k)}\right)\right\} \qquad [3]$$

of hypotheses h_i representing collections of target tracks into alternative scenes, and relative probability p_i of each hypothesis such that

$$\sum_{i=1}^{N_h(k)} p_i = 1 \qquad [4]$$

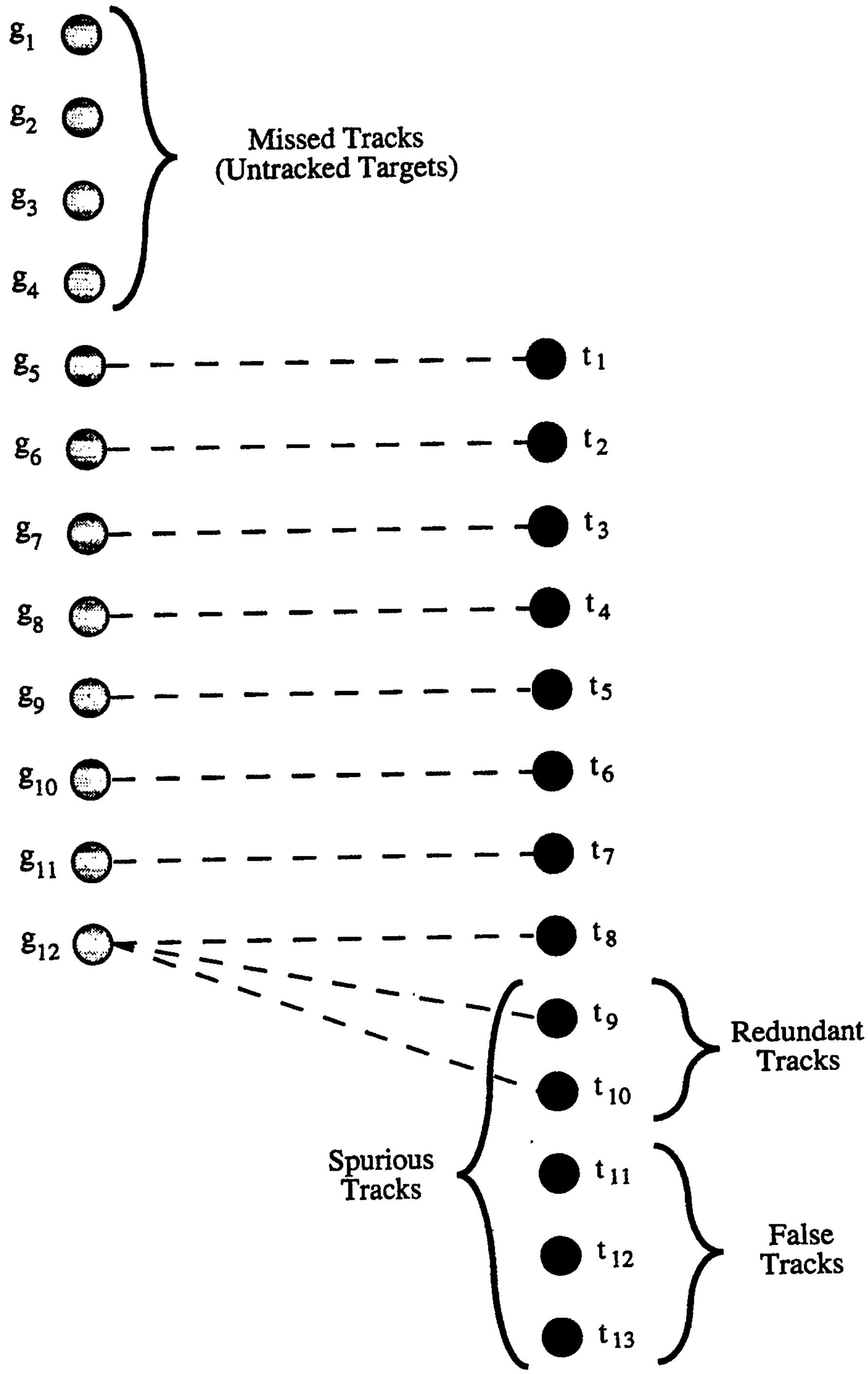

FIGURE 5 - Ground truth (g) and estimated (t) pictures (hypothesis "*h*", frame "*k*")

i.e., p_i is the probability that the hypothesis h_i is true (i.e., h_i is a correct assessment of ground truth). In turn, a hypothesis is assumed to be a list $h_i = \{t_1, t_2, ..., t_{N_t[h_i]}\}$ containing $N_t[h_i]$ tracks (Ref. 7).

A tracker output may therefore be interpreted as a weighted collection of estimates of ground truth which are assumed to span the available probability. That is, the tracker asserts that there are no other valid estimates of ground truth than the ones supplied in the tracker output (Ref. 10).

2.5.1 Individual Track Quality

Figures 6 and 7 illustrate some examples, regrouped with respect to the major MSDF subprocesses, of potential tracking problems affecting track quality. For example, a tracking algorithm may be numerically unstable (Fig. 6), causing the resulting track to be inaccurate and subject to track loss. Moreover, any mismatch, during a maneuver, between the real kinematic behavior of a target and the motion model assumed by a simple target state estimation algorithm can completely degrade the performance of the estimation technique.

As shown in Fig. 7 (a), track instability and track loss may also be caused by incorrect measurement associations with tracks. Moreover, track switch may occur when two or more targets are in close proximity (due to close parallel paths, or during a target encounter resulting from crossing paths). In such a situation, the measurements from other nearby objects can potentially be incorrectly correlated with the track on a given target.

Figure 7 (b) illustrates various cases of late track initiation and premature track deletion (or loss of track) which are two major track management issues.

2.5.2 Estimated Tactical Picture Quality

Often, as evidenced in Figs. 5 and 7 (c), there are divergences between the estimated tactical picture and the ground truth tactical picture (i.e., besides the individual track quality issues discussed above). In particular, there is no a priori correlation between the number of true targets N_g and the number of tracks N_t. In order to decide how well the MSDF system output O describes ground truth G, one must decide how correctly

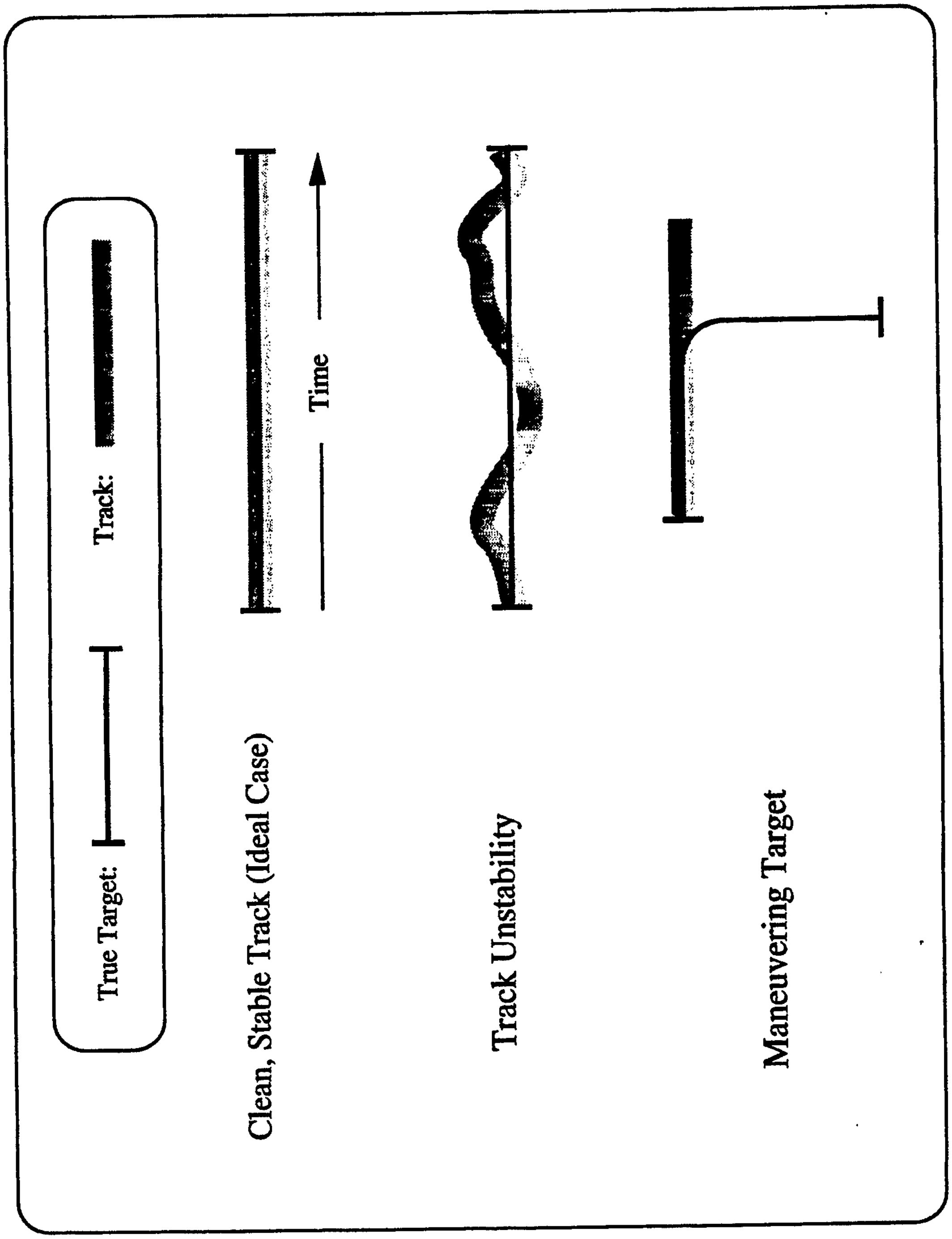

FIGURE 6 - Examples of potential target state estimation problems

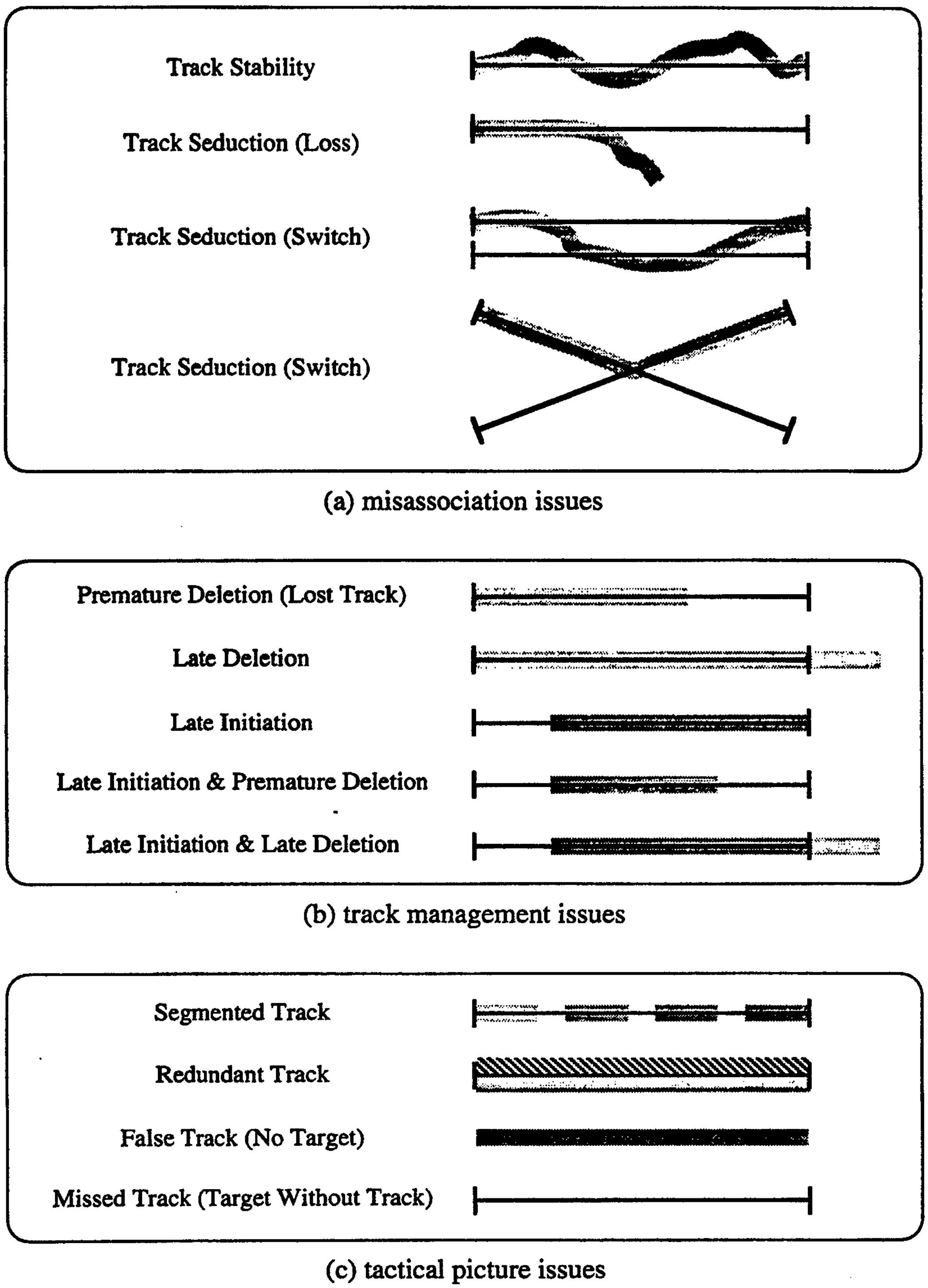

FIGURE 7 - Examples of other potential tracking problems

O has associated ground truth targets with the tracks in each hypothesis. In general, for each hypothesis h there will be four possibilities for associations between ground truth targets and its tracks (Ref. 7):

1) The track t in h uniquely corresponds to some ground truth target $g \in G$.
2) The track t in h corresponds to some ground truth target $g \in G$, but there is another track t' that was established prior to the track t, and that corresponds to the same ground truth target $g \in G$. In such a case, t is a redundant track.
3) The track t in h is false and does not correspond to any ground truth target.
4) Some ground truth targets have been missed (i.e., they are represented by no tracks in h at all).

For an MSDF system that tends toward a high incidence of redundant and/or false tracks, N_t may be much larger than N_g (i.e., the MSDF system tends to overcount). It is just as possible, however, that a tracking filter of the system (which is understood to include any logic which acquires new object tracks via track splitting, for instance) may be sufficiently "detuned" that it undercounts the number of objects in its perception of the target environment, so that N_t may be less than N_g (Ref. 5).

In scenarios where the stability of tracks is very poor (due to maneuvering targets or misassociation issues for example), multiple tracks may be consecutively initiated (and rapidly dropped) on the same target. Figure 7 (c) illustrates this situation as a segmented track. In such a case, the number of tracks N_t over the entire time interval of interest may also be much larger than the number of true objects.

2.5.3 Tracked-Object Taxonomy

A track classification (or taxonomy) that is used to identify the various types of mismatches between the estimated and ground truth pictures, and which facilitates performance evaluation for MSDF systems, is derived in Ref. 5. A slightly modified classification is described below. Because of various ambiguities, however, as discussed in more depth in Chap. 3, it may not be easy to determine to which taxonomic class any particular track belongs.

2.5.3.1 Clean Tracks

A clean, stable track is of the highest quality. It corresponds to a single object of interest in the environment over the track's entire history, without any pathologies such as misassociation (except, perhaps, during its initialization stage, when some bit of "bouncing around" may occur) or premature loss of track (Ref. 5).

With respect to misassociation, however, the rigor of this definition of a clean track may be relaxed by making use of the concept of track purity (Ref. 11). That is, a track may be classified as clean if its computed track purity score is above a pre-defined threshold.

2.5.3.2 Spurious Tracks

Spurious tracks include redundant (i.e., more than one track for one target), false (i.e., tracks for no targets whatsoever), and lost tracks. These are further discussed below.

A *redundant track* has the same characteristics as a clean track with the exception that a redundant track is subordinate to some other established track by virtue of its correspondence to the same physical object after a later initialization, i.e., of two tracks established on the same object, the later one will be considered redundant (Ref. 5).

A *false track* is of the poorest quality. It does not correspond to a physical object over any extended interval in its history, but instead "wanders off" immediately. This phenomenon can occur when the tracking system acquires what appears to be a real object, and resources are then scheduled to illuminate or scan the observation cell in which the target is expected to be at the next sampling instance. If the target is not in the expected observation cell (because there is no real object), the MSDF system may continue to extrapolate the object track until other logic determines that the track is spurious, and purges it from the system (Ref. 5).

Lost tracks correspond to once valid tracks that, for some reasons such as state estimation instability or miscorrelations, become spurious. As for the redundant and false tracks, a logic must determine at one point that the track is spurious, and purge it from the system.

2.5.3.3 Missed Tracks

Missed tracks are targets without tracks. Any tracker must, when confronted with a new sensor report, decide whether to associate it with existing tracks or else declare the existence of a completely new track. Expressed in different words, the MSDF algorithm must, whether explicitly or otherwise, have a way of estimating the likelihood that additional targets have not yet been detected (Ref. 7).

In its most general form, this likelihood is a number $r[h]$ associated with each hypothesis h which represents an estimate of the number of ground truth targets that have been missed by this hypothesis. At its simplest, it is a single number $L = r[h]$, independent of h, which represents an estimate valid for all hypotheses. The larger the value of L, the more likely it is that a new sensor report will be associated with a new track. Thus, the more intelligently L is estimated by the MSDF algorithm, the more efficiently it will generate hypotheses which accurately represent ground truth (Ref. 7).

The value of L is a piece of information in its own right with potential tactical usefulness. For example, consider the situation in which it must be decided whether or not a region is clear of targets and therefore "safe". If L is small, then the probability that the answer is "yes" is high; otherwise, it is low.

2.5.3.4 Dropped Tracks

A dropped track is one which does not exist at the end of the tracking interval (Ref. 5). A binary dropped track occurs when track quality has diminished below a certain threshold (e.g., as defined by the trace of its covariance matrix or the number of successive scans on which no observation data were received). A redundant dropped track occurs when the track is judged to be superfluous on the basis of a specific algorithm which detects the presence of redundant object tracks.

3.0 MSDF PERFORMANCE EVALUATION CONCEPT AND ISSUES

This chapter addresses some of the performance evaluation requirements and issues encountered in the use of MSDF systems in defence applications. In particular, this chapter discusses some of the reasons why MSDF systems are hard to evaluate. It also discusses the prediction of tracking performance using analytical evaluation methods. It very briefly addresses covariance analysis for use in design and evaluation of the tracking filters, expressions for use in the preliminary determination of observation-to-track correlation performance, and finally, Markov chain techniques for use in determining target track maintenance and false track statistics. The chapter ends with discussions on the importance of Monte Carlo simulation techniques, and on experimental trials for performance evaluation.

3.1 Performance Evaluation Requirements

Figure 1 illustrates the main components of the MSDF performance evaluation concept discussed in this document. The performance evaluation component implements the mechanism that allows the analysis of the absolute and relative performances of MSDF systems (in terms of tracking, identification, and global criteria), through comparisons of the measured and estimated tactical pictures with the ground truth picture.

There are several situations where objectively evaluating MSDF systems performance is important. First, the MSDF system designer typically has numerous choices when selecting a particular MSDF design concept for a given application. Making an intelligent selection demands the development of an approach for evaluating and comparing the performance of candidate systems. In particular, the functions of any element of a complex system must be understood, modeled, and quantitatively evaluated to determine the contribution that it provides to the effectiveness of the overall system. Some examples of the important system-level issues frequently raised about MSDF systems are:

- What is the best combination of sensors to meet a given set of detection probability, target discrimination, and target location requirements?
- What level of detection, discrimination, and location performance can be achieved by fusing a given set of existing sensors? What improvements

are accrued by adding sensors or improving the performance of individual sensors?

Second, the evaluation process will also be used to refine the performance of the selected MSDF system concept. The MSDF designer has the mandate to find ways to use the available sensing and processing resources as efficiently as possible. Hence, the evaluation process must allow the measurement of performance criteria for some targeted MSDF algorithms, providing quantitative inputs to support their optimization (Refs. 3, 6 and 12).

Finally, the quantitative performance assessment process is often applied to the MSDF elements of C^2 systems to determine the relative contribution that fusion provides to the military effectiveness of those systems (Ref. 4). As an example, the final goal of the MSDF project at DREV is to develop an optimal MSDF function that will enhance the tactical performance of the Command and Control System (CCS) of the ship. The performance evaluation capability is being developed to quantify these enhancements for a specific set of MSDF algorithms, proving the benefits of the MSDF function. The issues of system performance and system effectiveness are thus keys to establishing first, how well an algorithm, technique, or collection of techniques perform, and second, the extent to which these techniques may be used to achieve success on an operational mission (Ref. 13).

The reader is cautioned, however, that the relationship between the MSDF function and its contribution to improved military value is complex (Refs. 4 and 14). Overall, it is difficult to answer the following fundamental question: To what extent does an implemented MSDF system support the mission of a user? Serious analysis must demonstrate an appropriate selection of the MSDF system modeling approach, a thorough understanding of all contributing factors and functional relationships, attention to details, an ability to relate results to real-world data to validate accuracy of model elements, and sound judgment in evaluating the meaning and extent of results (Ref. 4). In this regard, Chap. 4 discusses in depth a methodology for MSDF performance evaluation.

3.2 Difficulties in MSDF Performance Evaluation

The difficulties in effective MSDF performance evaluation put both developers and users at a tremendous disadvantage, and represent a major obstacle to effective MSDF technology transition (Ref. 7). These difficulties hinder:

- engineering trade-offs,
- preparation of procurement specifications,
- conduct of the competitive selection process, and
- definition of meaningful acceptance tests.

Moreover, given these difficulties, performance evaluation may require a degree of sophistication that exceeds that of the MSDF algorithms themselves.

3.2.1 Lack of Global Measures Of Merits

The currently available Measures Of Merit (MOMs) are insufficient (Ref. 7). New MOMs that emphasize global performance must be invented. This is particularly true for multi-hypothesis MSDF algorithms that are expected to estimate the likelihood of alternative interpretations of the data. Because they produce not one but many answers, it has been difficult to establish a global concept of accuracy for them. They present a formidable challenge when it comes to engineering evaluation, competitive selection, and acceptance testing. Traditional MOMs, developed long ago for single-sensor single-hypothesis applications, do not adequately solve the issue of matching up a constructed track picture with the ground truth picture. They do not deal with the added value of multiple hypotheses.

3.2.2 Diversity of Aspects Involved

The evaluation of MSDF algorithms is complex because of the diversity of aspects involved (Ref. 7). MSDF systems provide answers to so many different questions. For any given dimension of a target estimate (e.g., location, velocity, attributes) a density function on estimate errors, basically a histogram, can be constructed. Such a density function gives an excellent global sense of the magnitude and dispersion of inaccuracies for that dimension. But what about other dimensions? What happens when an algorithm performs well in one regime, perhaps at the expense of another? It is difficult to establish the relative importance of answers, and there is a problem of adding apples and oranges to arrive at an overall assessment.

3.2.3 Complexity and Stochastic Behavior

The target estimates are a blend of random distributions, hard to evaluate because they jump around so much (Ref. 7). Even if we are able to repeat the same scenario exactly many times over, each repetition may be expected to produce a different set of answers, due to the natural distribution of sensor errors. When we attempt to draw conclusions about such systems we must cope with such behavioral dispersions. In order to evaluate them, at any confidence interval, we must take our evaluations over appropriately large sample sizes. Prudence also dictates that they be evaluated with suites of test conditions that are both diversified, and numerous enough to swamp out erratic statistical excursions. Such testing should at the same time be stressful enough to unearth pathological behaviors.

3.2.4 Location Effects

The accuracies that the MSDF systems achieve are seriously affected by location effects (Ref. 7). Sensor accuracies are always range dependent, and sometimes angle dependent. Therefore, target locations and kinematics at long ranges and poor azimuth or elevation will be less accurate. Environmental refraction can make targets in various range belts difficult or impossible to detect. Maneuvers that bring targets into close proximity or into the shadow of one another create correlation problems leading to reduction of accuracy. Because of these sundry location effects, the apparent accuracy of an MSDF system can change radically from one scenario to another. It is therefore very hard to establish an accuracy expectation for MSDF systems, except with reference to very explicitly defined scenarios.

3.2.5 Target Priorities

Some targets are more important than others (Ref. 7). If the tracker's behavior is to be governed by priorities, as we expect it to be, then the MOMs must ascertain how well the tracker is responding to these priorities. An example of high priority is an unknown target approaching at high speed. For such cases we may wish to focus our MOMs on a tracker's response time and accuracy against only one track. In other applications, we may wish to focus on a single class of targets. When the tracking system establishes such priorities, they must somehow be included in the evaluation environment.

3.2.6 Model Matching Issues

Sensor data fusion algorithms may incorporate various models that enable them to predict trajectories and behaviors, to infer from evidence, and to exploit negative information. These models are of two types: math models of physical behavior, and enumeration models that list attributes of objects. Simulators must include all these models in order to generate the synthetic reality that is used to stimulate and test an algorithm. In the operational setting, an algorithm's accuracy depends on how well its models agree with current reality. In a lab evaluation, the assessment of an algorithm's accuracy will depend on how well its models agree with the ones in the simulator. The models in the algorithm could very well be better approximations of reality than those in the simulator. But if they are different, the algorithm will be penalized for it in an evaluation. From this, it follows that the stimulator should have a higher fidelity than the algorithm being tested (Ref. 7). The details of the stimulator should be fully disclosed, so that accidental differences in modeling may be discovered and disqualification on the basis of moot distinctions avoided. Disclosure may also help uncover problems with the stimulator, whose correctness is very crucial yet never absolutely guaranteed.

3.2.7 Model Fidelity and Efficiency Issues

By definition, all models are an approximation of reality, and as a general rule the closer we get to reality the more expensive it becomes to gain another inch or so. Two questions must be asked: how good does the approximation must be? How much can we afford to pay? In an ideal tracking system, very high fidelity models would enable very accurate tracking. But, in tactical systems, the tracker is often called upon to respond to high data rates and time-critical requirements under severe limitations of size, weight, and power. These constraints militate against high fidelity and force the designer into many compromises. Recognizing these circumstances, the performance evaluation environment should insist that tactical algorithms be evaluated "at speed". It does no good to certify an algorithm's accuracy in the lab, only to discover in the field that it has to be emasculated to meet time and space constraints (Ref. 7).

3.2.8 Differences Between Algorithms

Differences, large and small, interfere with comparative evaluations (Ref. 7). Miscellaneous differences between algorithms make them difficult to compare. These differences result from the wide variety of target suites, sensors, and environments for which the techniques and algorithms have been designed.

3.2.9 Performance Evaluation Ambiguities

In complex tracking scenarios involving multiple targets and multiple sensors, the performance evaluation methodology must handle the ambiguities that create confusion about which target goes with a track (Refs. 15-18).

Evaluating the performance of tracking algorithms is usually straightforward for a simulation with a single target and one computed track, or in an environment of few, widely spaced targets and no false alarm or clutter. In such a sparse environment, a track is consistently updated with measurements from the same target. Track purity is thus ensured and the association of track-to-truth unambiguous.

However, multiple closely spaced targets and multiple tracks generate ambiguities that must be resolved during performance evaluation. Such a multiple target tracking environment typically involves many impure tracks. That is, a track is not consistently updated with measurements from the same target because some sensor observations of other targets, clutter or false alarms will be incorrectly associated with the track, and some sensor observations associated with the track will be unresolved closely spaced objects (Ref. 16), thus confusing which track is to be compared with the ground truth state of a target.

Furthermore, in a dense target environment, there may be missed tracks, redundant tracks, lost tracks, and false tracks. This raises the question of how to evaluate various measures of performance (e.g., state estimation error) with spurious or missing tracks (i.e., when the estimated number of targets is not the same as the actual number of targets). Figure 8 illustrates the ambiguities that may arise, for instance, when evaluating the state estimation performance (Ref. 18). A complete discussion of the MSDF performance evaluation ambiguities is out of the scope of this report. However, this issue is covered in depth in a separate document (Ref. 11).

ESTIMATED NUMBER OF TARGETS		ACTUAL NUMBER OF TARGETS: 0	1	2
	0	NO PROBLEM	MISSED TRACK ?	MISSED TRACKS ?
	1	FALSE TRACK ?	NO PROBLEM (MAY BE A FALSE TRACK)	AT LEAST ONE MISSED TRACK ?
	2	FALSE TRACKS ?	AT LEAST ONE FALSE TRACK ?	WHICH ESTIMATE GOES WITH WHICH ?

? ≙ WHAT IS THE VALUE USED FOR THE STATE ESTIMATION ERROR?

FIGURE 8 - Example of potential ambiguities when evaluating the state estimation performance (Ref. 18)

3.3 Analytical Evaluation Methods

At a certain state of the development, serious MSDF system design and evaluation typically require the development of a detailed Monte Carlo type simulation using complex models and algorithms. However, although it is an invaluable tool for coping with MSDF complexity, extensive Monte Carlo testing is both costly and time-consuming. Hence, it is usually convenient (and is an aid to the initial design process) to separately evaluate the MSDF elements, using less elaborate techniques, before they are included in the detailed simulation for final validation.

A problem of interest is thus the performance evaluation of the MSDF algorithms without recourse to extensive simulations (Refs. 6, 8, 19-25). In particular, in order to guide MSDF system design and preliminary evaluation, the analytic prediction of tracking performance may be desirable to estimate upper bounds for the performance given basic threat/sensor parameters, and to identify the key factors contributing to the tracking performance (Refs. 20-21). These bounds are typically based on the information content of the data, and are not related to any particular estimation technique. Based on the kind and amount of data involved, they provides an objective yardstick by which tracking and estimation concepts can be evaluated. No estimation technique can provide better accuracy, but there is no guarantee that the bounds can be achieved (Ref. 6).

If the upper bound associated with a specific concept does not meet the system requirements, the concept is unfeasible. If the upper bound does meet the system requirements, no firm conclusion can be made about feasibility. If the bound exceeds the requirements by far, the concept merits further attention and is probably a low-risk approach. If the bound barely meets requirements, it indicates that the concept is risky, and alternatives should be investigated.

Using analytical techniques, the sensitivity of the performance bounds to variations in each model (such as target maneuver model, encounter geometry, update intervals, sensor types and measurement accuracies) can be assessed directly, additional error sources (such as navigation errors in multiple platform data fusion) can be incorporated, and the effects of different fusion architectures can be explored. However, the optimistic nature of these bounds must be kept in mind. A detailed simulation is still required for an accurate assessment of a particular concept (Ref. 6).

3.3.1 Covariance Analysis

The value of Kalman filters for target tracking is well established, partially because many factors are easily taken into account with minimal effort. Update intervals, sensor types, measurement accuracies, and target maneuver models all affect the gains (or weights) applied to data as it is processed, and if the models are accurate, the result is an optimal estimate. Time variation of any of the models is also immediately accommodated, while still maintaining the estimate optimality. An evaluation of the accuracy of this estimate is available directly as the error covariance sequence computed as part of the filtering process. In linear filters, computing this error covariance is even independent of measurements or state estimates. For these and many other reasons, Kalman filter covariance analysis has become a very convenient, standard tool in evaluating tracking filter design and performance (Ref. 6).

As one application of covariance analysis, it is possible to examine the tracking process to evaluate the effects of mismatch between the Kalman filter model and the true target maneuver model in the actual system. The results will be used to determine the sensitivity to mismatches in the filter model (Ref. 8).

Covariance analysis can also be used to examine the effects of reducing the number of states in the Kalman filter. In particular, the third state, acceleration, can be removed. Thus, the process noise (or forcing) covariance matrix of the resulting two-state filter must be chosen to compensate for the missing acceleration state. In this case, covariance analysis is used to determine the degradation associated with the simplified two-state filter (as compared with the full three-state filter) (Ref. 8).

Finally, the effects of varying the sampling interval and the relative performance of various fixed-gain filters, as compared with a Kalman filter, can be readily computed using covariance analysis (Ref. 8).

The use of covariance analysis for navigation system performance analysis is a well-established procedure (Ref. 6). Equivalently, the use of covariance analysis for performance evaluation of multiple-sensor systems is not new.

3.3.2 Correlation Performance

For a preliminary design, the analytical (or theoretical) probabilities of correct correlation (P_{CC}), false correlation (P_{FC}), and correct decision (P_{CD}) are useful, important measures of predicted system performance (Ref. 8). These quantities (typically designed considering a single established track and one or more potential correlating observations) can be used to determine system feasibility and to aid in determining design parameters such as threshold setting.

The probability of correct association P_{CC} is defined as the probability for each track of being associated with the correct measurement. Its value depends on the object density in the measurement space, and on the average innovation standard deviation (which, in turn, depends on the measurement and target state estimate (in this case, the prediction) accuracy) (Ref. 20). Analytically, the probability of correct correlation (P_{CC}) is given by:

$$P_{CC} = P_D \, P_{CC/D} \qquad [5]$$

where P_D is the probability of detection of the target, and $P_{CC/D}$ is the probability of correct correlation given that a detection occurred. As a specific example, the latter is defined, for a nearest-neighbor correlation algorithm, to be the probability that the true return is inside the gate and that no false return has smaller normalized distance.

The probability of correct decision (P_{CD}) is P_{CC} plus the joint probability that no target detection occurs and that no correlation occurs given that there is no detection. Thus,

$$P_{CD} = P_{CC} + (1 - P_D) \, P_{NE} \qquad [6]$$

where P_{NE} is the probability that no extraneous sources fall within the gate (so that no correlation occurs) (Ref. 8).

These quantities (once their expressions are fully defined and expanded) can be readily computed from basic sensor and MSDF parameters, and then used to determine the feasibility of the design.

3.3.3 Markov Chain Analysis

Markov chain techniques are very convenient for deriving a variety of statistics related to track life for valid targets. In particular, an important measure of track retention performance is the probability of having a track that will not later be deleted. This is the cumulative probability of track confirmation without later deletion. This criterion is useful because it incorporates both search and track update capability into a comprehensive measure of tracking performance (Ref. 8).

The Markov chain approach can also be used to estimate statistics related to false tracks, such as the formation rate and density.

3.4 Monte Carlo Simulations

The analytic methods discussed above cannot be used to completely evaluate the effects of such complex processes as target evasive maneuvers, complex and correlated measurement noise and clutter processes, and miscorrelation between closely spaced target tracks. Given such complexity, classical covariance analysis and other simplified analytical methods will not be adequate to even begin to perform MSDF algorithm evaluation, and can indeed be very misleading. Hence, before a system can actually be implemented, it is usually necessary to evaluate its performance through the Monte Carlo simulation techniques (Ref. 8).

Performance evaluation under laboratory conditions, where targets, environment, and sensors are only simulated, is also a necessary preliminary to operational testing. It provides the highly flexible and repetitive testing capabilities that, because of time and very high cost, could never be duplicated in the field (Ref. 7). Moreover, it also establishes a high degree of confidence in the MSDF system before that system is exposed to uncertain circumstances in the field. With Monte Carlo simulations one is able to:

- explore the limitations, find out where breakage points are, where worst case occurs,
- create artificial, unrealistic situations, that reveal unsuspected, perhaps faulty behaviors,

- gather enough data to establish the shape and magnitude of error density plots,
- determine whether performance is balanced across a spectrum of conditions,
- gain assurance that design and coding errors have been eradicated,
- gather enough data to distinguish the effect of design changes.

The investigation of the Multiple Hypothesis Tracking (MHT) algorithm (Ref. 9) is one of the main drivers toward the use of Monte Carlo simulations. Evaluating the performance of MHT algorithms is important in predicting the performance of surveillance and tracking systems. Because of its near optimality (although any implementable version of this algorithm is often only suboptimal), the performance of the tracking algorithm under various conditions should provide a reliable estimate of what a sensor system is capable of when using the best tracking algorithm (Ref. 19). At this moment, Monte Carlo analysis is the only known way of predicting the performance of MHT algorithms if one needs accurate prediction of tracking performance in a realistic, and thus complex, environment.

Given its importance, the emphasis is given for the MSDF project at DREV (and within this document) to MSDF performance evaluation through Monte Carlo computer simulations. However, a complete discussion of some important implementation details and issues is deferred to a subsequent document (Ref. 11).

3.5 Experimental Trials

Testing of tracking algorithms and systems in computer simulations is but one step in the process of development test and evaluation. After successful simulations, verification and validation ultimately need to be accomplished in the operational environment, with actual targets, environmental phenomena, and sensors. However, an advantage to computer simulations is that the ground truth data that are needed for performance evaluation are readily available unlike in real-world tests where precise truth data can be very hard to obtain (Ref. 15). It is also very expensive in the operational environment to create the number and variety of test conditions that are required to validate an algorithm's correlational and statistical inferencing capacities (Ref. 7).

4.0 PERFORMANCE EVALUATION METHODOLOGY

As discussed in the previous chapter, evaluating MSDF systems is a complex and difficult activity. Haphazard approaches run the risk of introducing biases into the evaluation or generating completely spurious results. Therefore, a well-defined methodology for evaluating these systems must be constructed (Ref. 3). Rigorous and methodical testing is the only way to guarantee that the selected algorithms will work in a military application, and be the best choice from among several candidates (Ref. 7).

The intent is not to discount subjective evaluations, but rather to provide objective measures to support these subjective evaluations. When using subjective measures of system performance, it is important that these considerations be kept separate whenever possible from the computations of the objective, numerical MOMs. Failure to maintain this separation may result in highly questionable system evaluation (Ref. 3).

The classical system analysis methodology may be applied to the analysis of MSDF systems to compare alternative approaches (sensors, algorithms, architectures, etc.) and estimate the relative merits of candidate systems. This chapter, mostly based on Refs. 4 and 13, suggests a framework for the definition and discussion of the MSDF performance evaluation process; this framework is summarized in Table II. Much of the rational for and many of the issues raised about this framework are derived from "good" system engineering concepts, and are intended to sensitize MSDF researchers to the need for formalized performance evaluation methods to quantify or otherwise evaluate the marginal contributions of the MSDF process to program/system goals.

The framework outlined in this chapter for evaluating and comparing alternative MSDF systems in a variety of common target environments consists of many facets. The performance evaluation activities will be characterized by the following seven essential steps:

- Establish the context
- Select a philosophy
- Define specific objectives
- Construct alternatives
- Select performance evaluation criteria

TABLE II
Performance evaluation methodology steps (Refs. 4 and 13)

1	**Establish the Context**	Overall Goals and Objectives • Program and Mission Goals • R&D Objectives: - Proof of Concept Demonstration - Feasibility Test - Prototype Development - Production Prototype
2	**Select a Philosophy**	Establish / Emphasize a Perspective • Black-Box vs White-Box • Organizational • Economic • Informal • Formal
3	**Define Specific Objectives**	Specify • Quantitative Information to be Obtained • Qualitative Information to be Obtained • Decisions to be Made • Scope of the Analysis • Independent and Dependent Variables • Basic Assumptions
4	**Construct Alternatives**	Specify • Candidate Systems • Candidate Architectures
5	**Select Performance Evaluation Criteria**	Hierarchy • Criteria: Function of One or Many Measures of Merit • Measures of Merit: MOP, MOE, MOFE (Function of Metrics) • Metrics: Dimensional Parameters (Must be Observable)
6	**Develop an Approach**	A) System Modeling Approach: • Inputs, Process Model, Outputs • Scenarios B) Procedures: • Metric Gathering / MOM Computation Strategy C) Experimental Design (Evaluation Setup): • Test Case Design: - Test Case Characteristics - Number of Test Cases - Sequence of Test Cases • Standards • Analytical Framework
7	**Analyze the Results**	Analyze to Determine • Uncertainties in the Outcomes • Significance of Marginal Differences • Output / Input Variables Relations • Measures of Merit Sensitivity • Candidate Recommendation

- Develop an approach
- Analyze the results

4.1 Specification of the Context

The assessment of a delivered value for defence systems must be made in light of system or program objectives. In the design and development of such systems, many "translations" of the stated objectives occur as a result of the systems' engineering process, which both analyzes (or decomposes) the goals into functional and performance requirements, and synthesizes system components intended to perform in accordance with these requirements. Throughout this process, however, the objectives must be kept in view because they establish the context in which the value will be judged (Ref. 13).

The context therefore reflects what the program (and the MSDF process or system within it) is trying to achieve. That is, the context refers to what the purposes of building the system at hand are. The goals are typically reflected in the program's "name", such as a "Proof of Concept" program or "Production Prototype" program, etc.

4.2 Selection of the Philosophy

There are also several "translations" of objectives which occur for the performance evaluation activities themselves. In particular, performance evaluation philosophies follow from statements about goals and objectives. Philosophies primarily establish or emphasize perspectives for performance evaluation that are consistent with, and can be traced to, the goals and objectives: they establish the purpose of "investing" in the performance evaluation process. Philosophies also provide guidelines for the development of performance evaluation criteria, for the definition of meaningful cases and conditions and, importantly, a sense of "satisfaction scale" for test results and value judgments which guides the overall investment of precious resources in the performance evaluation process. That is, performance evaluation philosophies, while generally stated in non-financial terms, do in fact establish economic philosophies for the commitment of funds and resources to the performance evaluation process (Ref. 13).

Establishing a philosophy for performance evaluation of an MSDF process is tightly coupled to the establishment of what the MSDF process boundaries are. In general, it can be argued that the performance evaluation of any process within a system should attempt the longest "extrapolation" possible in relating process behavior to program goals, i.e., the evaluation should endeavor to relate process test results to program goals to the extent possible. This entails first understanding the MSDF process boundary, and then assessing the degree to which MSDF process results can be related to superordinate processes. For defence systems, this means assessing the degree to which MSDF results can be related to mission goals.

The simplest example of this "philosophy" notion is reflected in the "black-box" or "white-box" viewpoints for performance evaluation, from which either external (i.e., I/O) or internal (i.e., procedure execution) behaviors are examined. There are some other philosophies that could be established, however, such as organizational (i.e., one examines the benefits of MSDF products accruing to the system-owning organization), economic (i.e., one explicitly focuses on some sense of economic value of the MSDF results (e.g., weight, power, volume, etc.)), informal (MSDF results are measured against some human results or expectations), and formal. The latter is the class of philosophies in which the evaluation is carried out according to appropriate formal techniques which prove or otherwise rigorously validate the program results or internal behaviors (e.g., proofs of correctness, formal logic test, formal evaluations of complexity). And finally, another point of view revolves about the research or development goals established for the program.

Philosophies aside, there are the "acid tests" which should always be conducted by the evaluator:

- results with and without fusion (multi-sensors versus single-sensor (best sensor?)),
- results as a function of number of sensors or sources (e.g., single-sensor, *2, 3, ..., N* sensor results for a common problem).

4.3 Definition of Specific Objectives

The specific objectives of the analysis must be clearly stated in terms that specify the quantitative and qualitative information to be obtained, decisions to be made as a result of analysis (and the criteria for decision-making), scope of the analysis, independent and dependent variables, and basic assumptions. Objectives of a typical MSDF analysis follow (Ref. 4):

Example: Compare and rank the target detection-location capabilities of candidate MSDF systems that may use a four-sensor suite, employing sensor types S1, S2, S3, S4, S5, or S6, and MSDF algorithms A1 or A2.

Determine the relative performance for both jamming and benign environments with jamming effects to equally degrade sensors S1 and S5 in detection capability. Consider four representative scenarios, which range from large homogeneous target distributions to small concentrated target clusters.

Assume that all sensors are collocated and that there is no delay between measurement time and report time for sensors S1, S2, S4, S5, and S6. Assume a 5 second delay between S3 measurement and report time.

4.4 Construction of Alternatives

Using the available system resources specified in the objectives, candidate data fusion system architectures are constructed. The resource variables (usually sensors, sensor performance levels, sensor locations, and processing techniques) usually provide a large number of possible combinations; and a subset of candidates is intelligently selected to represent the primary categories of architectural alternatives (Ref. 4).

Example: For the preceding example, a set of system candidates may be defined by the following combinations of sensors and data fusion algorithms:

System Candidate A: S1, S2, S3, S4, A1
System Candidate B: S1, S3, S5, A2
System Candidate C: S1, S5, S6, A2

4.5 Selection of Performance Evaluation Criteria

Once having embraced one or another of the philosophies, there exists a perspective from which to select a set of criteria according to which the quality and correctness of the performance evaluation results or inferences will be judged. Obviously, the criteria for performance evaluation must also be selected considering the actual available data and the particular system implementation (Ref. 12). These various criteria will collectively provide a basis for evaluation. It is important at this step to realize the full meaning and subsequent relationships impacted by the selection of such criteria (Ref. 13).

There should be a functionally complete hierarchy which emanates from a criterion as follows:

- Criterion: A standard, rule, or test upon which a judgment or decision can be made.
- Measures of Merit: The "dimensions" of a criterion (i.e., the factors into which a criterion can be divided) through which judgments on the criterion can be made.
- Metrics: Those attributes of the MSDF process or its parameters or processing results which are considered easily and straightforwardly quantifiable or able to be defined categorically, and, importantly, are observable (i.e., which can be measured during performance evaluation experiments). Metrics are also often called "dimensional parameters".

Hence, in the most general case, there is a functional relationship as:

$$\begin{aligned} \mathit{Criterion} = \mathit{fct}\,[\,&(\mathit{Measure}_i = \mathit{fct}\,(\mathit{Metric}_j, \mathit{Metric}_k, \ldots), \\ &(\mathit{Measure}_m = \mathit{fct}\,(\mathit{Metric}_n, \mathit{Metric}_p, \ldots), \\ &\mathit{etc.},\,] \end{aligned} \qquad [7]$$

One reason to establish these relationships is to provide for traceability of the logic applied in the performance evaluation process. Another rational which argues for the establishment of these relationships is in part derived from the requirement to at least estimate predicted system behaviors against which to compare actual results. Such prediction must occur at the metric level; predicted and actual metrics subsequently form the basis for comparison and evaluation. The prediction process must be functionally consistent with this hierarchy. Failure to do so may in fact invalidate the overall approach to the performance evaluation process (Ref. 13).

Each metric, measure, and criterion also has a scale which must be considered. Moreover, the scales are often incongruent so that some type of normalized "figure of merit" approach may be necessary in order to integrate metrics on disparate scales and construct a unified, quantitative parameter for making judgments.

Representative criteria focused on the MSDF process per se are listed below for the numerically dominated Level-1 MSDF process (position and identity (Ref. 4)), and the symbolic oriented Level-2 and 3 MSDF processes (situation and threat assessment):

- Correctness in reasoning
- Repeatability / consistency / reliability
- Computational complexity
- Time-critical performance
- Human / computer interaction
- Ease of implementation
- Pertinence
- Contribution to organizational goals
- Quality or accuracy of decisions / advice / recommendations
- Intelligent behavior
- Robustness
- Computational efficiency
- Ability to update
- Ease of use
- Cost (effectiveness / total cost)
- Trustworthiness
- Adaptability

4.5.1 Measures of Merit

Eventually, the criteria for evaluating the various alternatives must be defined quantitatively in the form of measures of merit that can be determined for each candidate. The measures must allow discrimination between alternatives and be appropriate to answer the questions stated in the analysis objectives (Ref. 4).

Example: For the preceding example, "target detection-location performance" could be quantified by the following standard measures:

- Detection probability
- False alarm rate
- Detection range as a function of the probability of detection
- Target location accuracy as a function of detection range

4.5.1.1 Hierarchy from MSDF Performance to Military Effectiveness

Because sensors and fusion are contributors to improved information accuracy, timeliness, and content, a major objective of many fusion analyses is to determine the effect of these contributions to military effectiveness. Ultimately, the performance of tracking algorithms is judged by the success, or lack thereof, of the mission they support. The destruction of a target by an interceptor guided, in part, by tracking information provides one vivid, obvious measure of success (Refs. 15-16). In this regard, most of the CPF CCS overall integration tests are designed to demonstrate that the tracking systems can successfully provide sufficient quality data to the fire control system, enabling it to select the proper fire control solutions and to complete them successfully for every type of warfare (AAW, AWW, etc.) (Ref. 26).

The military effectiveness must be quantified, and numerous quantifiable measures of merit can be envisioned: engagement outcomes, exchange ratios (the ratio of blue-red targets killed), total targets serviced, and so on. The ability to relate MSDF performance to military effectiveness is difficult because of the many factors that relate improved information to improved combat effectiveness and the uncertainty in modeling them (Ref. 4). These factors include:

- cumulative effects of measurement errors that result in targeting errors,
- relations between marginal improvements in data and improvements in human decision-making, and
- effect of improved threat assessment on survivability of own forces.

These factors and the hierarchy of relationships between MSDF performance and military effectiveness must be properly understood to develop measures and models that relate them. The Military Operations Research Society (Ref. 27) has recommended a hierarchy of measures, shown in Table III, that relate performance characteristics of C3 systems (including MSDF) to military effectiveness. This hierarchy of measures includes dimensional parameters, measures of performance, measures of effectiveness, and, for tactical MSDF systems, measures of force effectiveness (Refs. 4 and 14).

Dimensional Parameters

Dimensional parameters (that are equivalent to the metrics previously introduced) are the first and most basic measure or quantification of a MSDF system. They seek to quantify the typical properties or characteristics inherent in the physical entities that directly define the elements of the MSDF system (sensors, processors, communication channels, etc.). The values of these parameters directly describe the behavior or structure

TABLE III

Four categories of measures of merit (Refs. 4, 13 and 27)

Measure	Definition	Typical Examples
Measures of Force Effectiveness (MOFE)	Measure of how a C3 system and the force (sensors, MSDF system, weapons, C3 system) of which it is part perform military missions.	Outcome of battle Cost of system Survivability Attrition rates Exchange ratio Weapons on targets
Measures of Effectiveness (MOE)	Measure of how a C3 system performs its functions within an operational environment.	Target nomination rate Timeliness of information Accuracy of information Warning time Target leakage Countermeasure immunity Communications survivability
Measures of Performance (MOP)	Measures closely related to dimensional parameters (both physical and structural) but measure attributes of system behavior.	Detection probability False alarm rate Location estimate accuracy Identification Probability Identification range Time from transmission to detect Communication time delay Sensor spatial coverage Target classification accuracy
Dimensional Parameters	The properties or characteristics inherent in the physical entities whose values determine system behavior and the structure in question, even when not operating.	Signal-to-noise ratio Bandwidth, frequency Operations per second Aperture dimensions Bit error rates Resolution Sample rate Antijamming margins Cost

of the MSDF system and should be considered to be typical measurable specification values (bandwidths, bit-error rates, physical dimensions, etc.). Hence, at the lowest level of evaluation, an MSDF system would be characterized by the extent to which an implemented system meets specified dimensional parameters.

Measures of Performance (MOPs)

MOPs are the measures that describe the important behavioral aspects of the MSDF system (Ref. 4). In a sense, MOPs characterize how well an implemented system of sensors and data fusion algorithms performs with respect to transforming signal energy either emitted or reflected from a target (i.e., an external physical situation) into an output representation via state vectors or identity declarations (resulting in a level-1 MSDF database). Typically, MSDF MOPs operate on tracking and on identification in terms of time efficiency and precision (Ref. 12).

MOPs involve not only sensors and system parameters, but also an interplay of those dimensional parameters with implemented techniques to perform MSDF (Ref. 14). Hence, MOPs are often functions of several dimensional parameters to quantify in a single variable a significant measure of operational performance. Intercept and detection probabilities, for example, are important MOPs that are functions of several dimensional parameters of both the MSDF system and the targets being detected.

The evaluation of tracking performance is not limited to the evaluation of state estimation and prediction errors. Practicality dictates that several different MOPs be evaluated, so that users with different interest can find the kind of answers they want (Ref. 7). Table IV (compiled from all the references) presents a set of objective numerical MOPs that can be used to assess the performance of the MSDF system from a variety of perspectives, and which are relevant to typical mission goals. These MOPs have been classified into six main categories:

- measurement quality
- individual track quality
- estimated tactical picture quality
- computational performance
- communication performance (not really MSDF MOPs)
- overall rating

TABLE IV

MSDF measures of performance

1	**Measurement Quality**	• Target Detection and Leakage - P_D, P_{FA} - Detection Range - Time in Volume of Interest Prior to Detection - Time from Emission to Detection • Kinematic Measurement Accuracy (Error Mean / Variance) • Attribute Measurement Accuracy - Bayesian Percentage Attribute Miss (BPAM) - Non-Bayesian Percentage Attribute Miss (NPAM) - Mean Number of Misidentified Attributes (MNMA) - Identification Range - Range of "Unknown" Identification - Range of "Correct" Identification - Range of "Positive" Identification - Time in Volume of Interest Prior to Identification • Number of Unresolved Closely Spaced Objects • Number of False Measurements • Average Track (Target) "Exposure" • Countermeasure Immunity
2	**Individual Track Quality**	• Kinematic Track Accuracy - Radial Miss Distance - State Estimation Error - State Estimation Bias - Credibility of Filter Calculated Covariance - Accuracy of Filter Calculated Covariance - Track Stability - Number of Tracking Filter Re-Initializations - Tracking Decision Confidence (Track Quality Index) • Identification Track Accuracy - Belief and Support Evolution Analysis - Bayesian Percentage Attribute Miss (BPAM) - Non-Bayesian Percentage Attribute Miss (NPAM) - Mean Number of Misidentified Attributes (MNMA) - Identification Range - Range of "Unknown" Identification - Range of "Correct" Identification - Range of "Positive" Identification - Correlation of ID Lists - Time in Volume of Interest Prior to Identification • Association Performance - Probability of Correct / Incorrect Validation - Probability of Correct / Incorrect Correlation - Probability of Correct / Incorrect Decision - Track Depth - Number of Track Switch - Strict Sense Track Purity - Loose Sense Track Purity - Modified Loose Sense Track Purity - Correct Assignment Ratio • Track Management Statistics - Composite (M/N) Track Initiation Logic - Track Acquisition (Confirmation) - Probability of Track Acquisition - Probability of Establishing a False Track - Expected Time to Firm - Track Deletion - Time Required to Delete a Track on a Target That Leaves theVOI (Late Track Deletion) - Number of Premature Track Deletions (Track Losses) • Threat Number (Correct / Incorrect, Timing) • Quick Reaction (Correct / Incorrect, Timing)

TABLE IV (contd)

3	**Estimated Tactical Picture Quality**	• Target Detection and Leakage - Number of Detected Targets in the Volume of Interest - Number of Undetected Targets in the Volume of Interest - Percent of Detected Targets in the Volume of Interest - Percent of Detected Targets at a Given Range - Percent of Targets that Evade Detection • Track Management - Range of Confirmation for Targets (Hostile vs All) - Raid Assessment - Probability of Having at Least N Confirmed Tracks - Probability of Having at Least N Confirmed Tracks That Will Not Later Be Deleted - Expected Number of Tracks - Number of Targets in Tracks - Number of Missed Tracks - Percent of Targets Correctly Confirmed (Coverage Degree) - Estimated Targets in Track / Actual Targets in Track - Time in Volume of Interest Prior to Raid Assessment - Kill Assessment - Percent of Targets Correctly Assessed as "Killed" - Percent of Targets Correctly Assessed as "Alive" - Average Time Required to Delete a Track on a Target That Leaves theVolume of Interest (Late Deletion) - Number of Dropped Tracks - Number of Premature Track Deletions • Association Performance - Expected Number of Correct / False Correlations Per Scan - Misassociation Index • False Tracks (Probability and Percentage) • Segmented Tracks - Tracking Time / Expected Track Length - Number of Initializations / Track Loss (Runaway) - Number of Re-Acquisition Performed • Redundant Tracks (Path Multiplicity) • Kinematic Accuracy Performance - Track Variance for Hostile Targets - Track Variance for Priority Targets - Track Variance for all Targets - Percent of Hostile Targets with Launch Quality Tracks • Target Identification Performance - Identification Range - Average Range of "Unknown" Identification - Range of "Correct" ID for Targets (Hostile vs All) - Range of "Positive" ID for Targets (Hostile vs All) - Percent of Targets Correctly Identified (Hostile vs All) - Correlation of ID Lists • Threat Number (Order and Timing in the Scenario) • Quick Reaction (Order and Timing in the Scenario)
4	**Computational Performance**	• Number of Hypotheses, Tracks, ID Propositions, etc. • Efficiency (Ratio of Stable Tracks to Total Tracks) • Ratio of Targets Tracked to System Tracking Capacity • Processing Burdens (MIPS, Memory, Bus Bandwidth Utilized)
5	**Communication Performance**	• Reporting Bandwidth • Timeliness of Information (Communication Time Delay) • Communication Reliability (Including Survivability)
6	**Overall Rating**	• Composite dBn Rating • A Single Measure of Effectiveness • Amount of Information

The MSDF MOPs listed in Table IV are closely tied to notions of physical reality, and can be applied both to comparative evaluations of MSDF system alternatives (in which relative performance is the issue), and to evaluations of MSDF system performance relative to what can be achieved in any given target object environment (i.e., relative to some kind of upper bound) (Ref. 5). Indeed, these MOPs allow performance evaluation across a multiplicity of target environments for a given MSDF system, or across a multiplicity of tracking systems for a given target environment or matrix of environments. In particular, the kind of target tracking scenarios that cause the MSDF system some degree of difficulty can be readily assessed with these performance measures.

The foundation and unifying concept in the metrical approach behind Table IV is the measurement of the distance between truth and conjecture, in the multi-dimensional space that one uses to describe target locations, kinematics, attributes and identity. The concept of distance in this multidimensional space, where the number of targets in ground truth is not necessarily the same as that in measured or estimated pictures, is not simple. In particular, one is forced to develop concepts of distance in attribute space. Nevertheless, distance and probabilities (across an ensemble of hypotheses) are the only reliable and consistent parameters for determination of quality indices (Ref. 7).

Details about each MSDF MOP and implementation issues are beyond the scope of this document. These aspects will be covered in depth in a separate document (Ref. 11).

Measures of Effectiveness (MOEs)

MOEs represent a hierarchical level above MOPs, and are more difficult to identify than MOPs, because MOEs seek to provide a measure of the ability of a fusion system to assist in completion of an operational mission (Ref. 13). MOEs gauge the degree to which a system function was successfully performed. For tactical military systems, MOE examples might include target nomination rate, warning time, target leakage, immunity to countermeasures, survivability, and other factors (Ref. 14).

Measures of Force Effectiveness (MOFEs)

These are the highest level measures that quantify the ability of the total military force (including the MSDF system) to fulfill its mission. Typical MOFEs are derived from battle outcomes: survivability, rates and ratios of attrition, exchange ratios, measures of weapon success, and functions of these variables (Ref. 14). However, in the overall mission definition, factors other than outcome of the conflict (e.g., cost, size of force, composition of force) may also be included in the MOFE.

4.6 Development of a Performance Evaluation Approach

Another element of the performance evaluation framework is called the *approach* element, through which tests and analyses that are consistent with the philosophy can be defined and conducted. By *approach* is thus meant a set of activities which are both procedural and analytical, and which generate both the results of interest (via analytical operations on the metrics), as well as decisions based on those results and in relation to the criteria. The approach consists of the following components (Ref. 13):

- a system modeling approach,
- a procedure, which is a metric gathering and MOP computation paradigm, and
- an experimental design (evaluation setup), which defines (1) the test cases, (2) the standards for evaluation, and (3) the analytical framework for assessing the results.

These components are briefly discussed below.

4.6.1 Development of an MSDF System Modeling Approach

The candidate systems must be modeled by defining inputs, process models and outputs (Ref. 4). In the MSDF project at DREV, this is addressed by using the CASE-ATTI (Concept Analysis and Simulation Environment for Automatic Target Tracking and Identification) system. CASE_ATTI is a highly modular, structured, and flexible simulation environment providing the algorithm-level test and replacement capability required to study and compare the technical feasibility, applicability and performance of advanced, state-of-the-art MSDF techniques. Figure 9 illustrates the global structure of

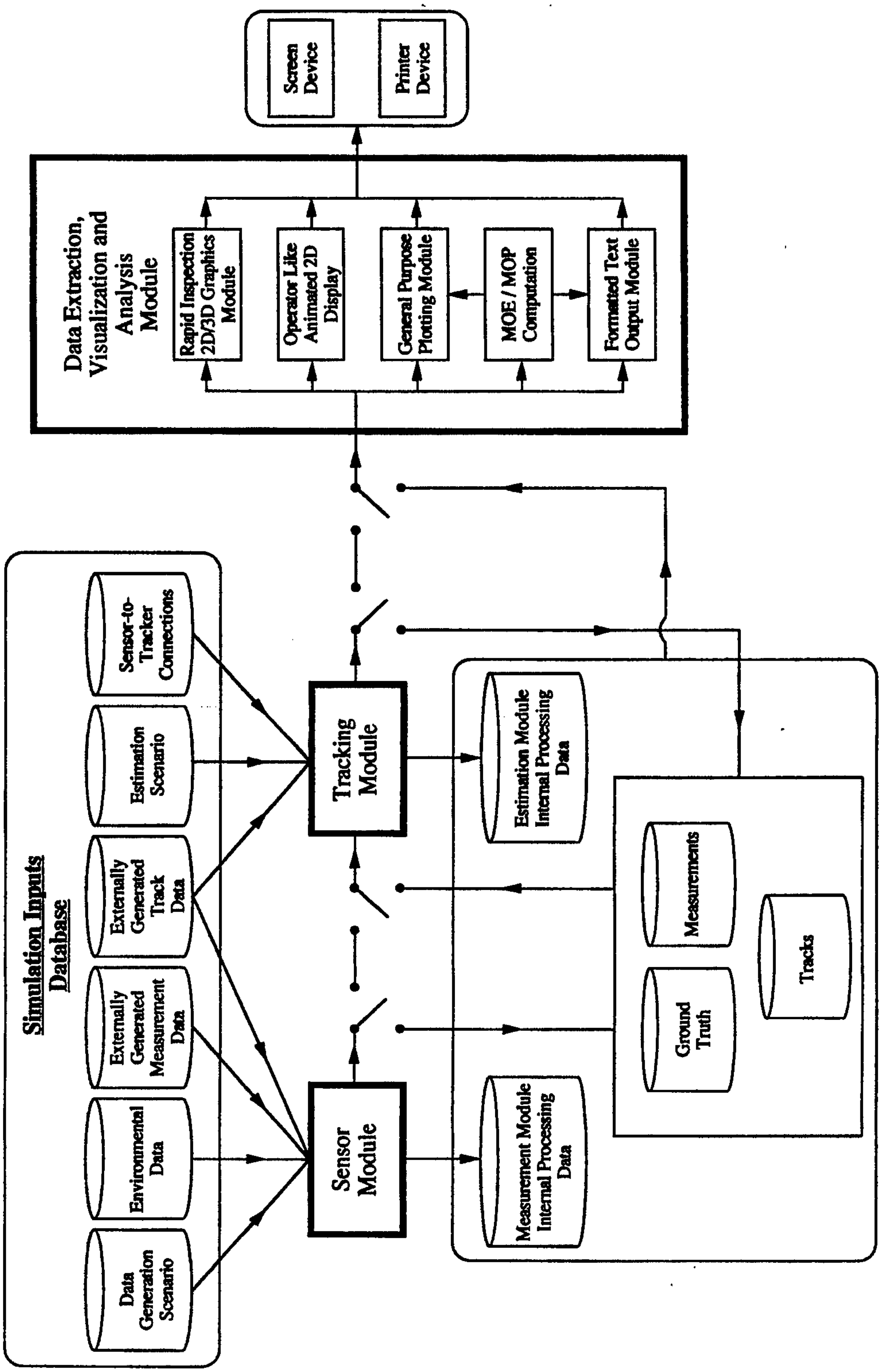

FIGURE 9 - Global structure of the CASE_ATTI system

the CASE_ATTI system. A basic description of this testbed was given in a previous document (Ref. 2). The enhancement of CASE_ATTI to fully support advanced MSDF performance evaluation activities will be described in a subsequent document (Ref. 11).

An important aspect of the modeling process is the development of scenarios representative of the real world. MSDF system performance (in both accuracy and speed regimes) is usually quite sensitive to the parameters of driver scenarios (Ref. 7). Indeed, varying scenario parameters such as the number and position of targets can drastically alter the performance of a system. Hence, prudence dictates that MSDF systems be evaluated over a set of scenarios that is both diversified and extensive enough to average out the statistical variations, while at the same time stressing enough to uncover pathological behaviors.

The scenario variables must include all critical entities, behaviors and events that influence the outcome of the analysis. Because these variables are usually numerous (scenario geometry, electronic environments, target quantities and mixes, physical environments affecting sensor and communication performance, etc.), the selection of a manageable set of truly representative scenarios is critical to the meaningfulness of the modeling process (Ref. 4). Hence, in order to provide a means to consistently evaluate MSDF systems performance, a set of standard benchmark scenarios should be defined. Unfortunately, such benchmark scenarios to test MSDF systems do not currently exist.

Since for any given scenario the "correct" target environment assessment may not be easily determined, the scenarios must be defined with explicit objectives and performance expectations in mind (Ref. 3). Examples of interesting scenarios are as follows:

- A random arrival scenario could be used to determine the dependence of the MSDF system detection, leakage, and tracking performance on target density (i.e., the number of targets in the current ownship volume of interest (or detection envelope)) and arrival rates.
- A closely spaced target pair offers a good test of the tendency of the MSDF system to misassociate measurements.
- A suddenly maneuvering target offers a good test of the ability of the tracking filter to cope with extraneous accelerations.

- A parallel formation of targets offers an excellent test of the track acquisition and association capabilities of the MSDF system.
- Loosely grouped objects offers a good test of the ability of the MSDF system to distinguish individual objects from a pack and sustain stable tracks without heavy redundancy.
- A large group of targets in a dense pack offers a stern test of the ability of the MSDF system to distinguish and sustain individual tracks in a crowded environment.

4.6.2 Performance Evaluation Procedure

The performance evaluation procedure is a metric gathering and MOP computation paradigm. At first sight, MOP computation may seem to be a rather straightforward process. However, as previously discussed in Section 3.1.9, there are fundamental ambiguities in evaluating multiple target tracking algorithms that are not encountered in performance evaluations with only a single object. These issues must be addressed before measures of performance, such as state estimation error and track purity, can be computed (Ref. 15).

A two stage process is proposed by Refs. 15 and 16 to resolve the ambiguities in performance evaluation of MSDF systems in a dense target environment. This two-stage method is illustrated in Fig. 10.

One first needs to relate tracks to targets (i.e., make the decision of which target to compare with a track) so that measures of performance can be computed. The first stage of the proposed performance evaluation procedure is thus the use of an assignment algorithm to assign tracks to true targets, based on some assignment criterion. Note that in evaluating performance, the spurious and missed tracks should be identified and counted in addition to resolving the ambiguities for the valid tracks. Hence, the output of the track-to-truth assignment stage includes not only the set of assigned target-track pairs, but also the set of unassigned targets (which are considered as missed tracks), and the set of unassigned tracks (which are treated as spurious tracks).

After track and truth have been associated in the first stage (i.e., for each assigned track-truth pair), various measures of performance for the main functions of an MSDF algorithm can be computed in the second stage.

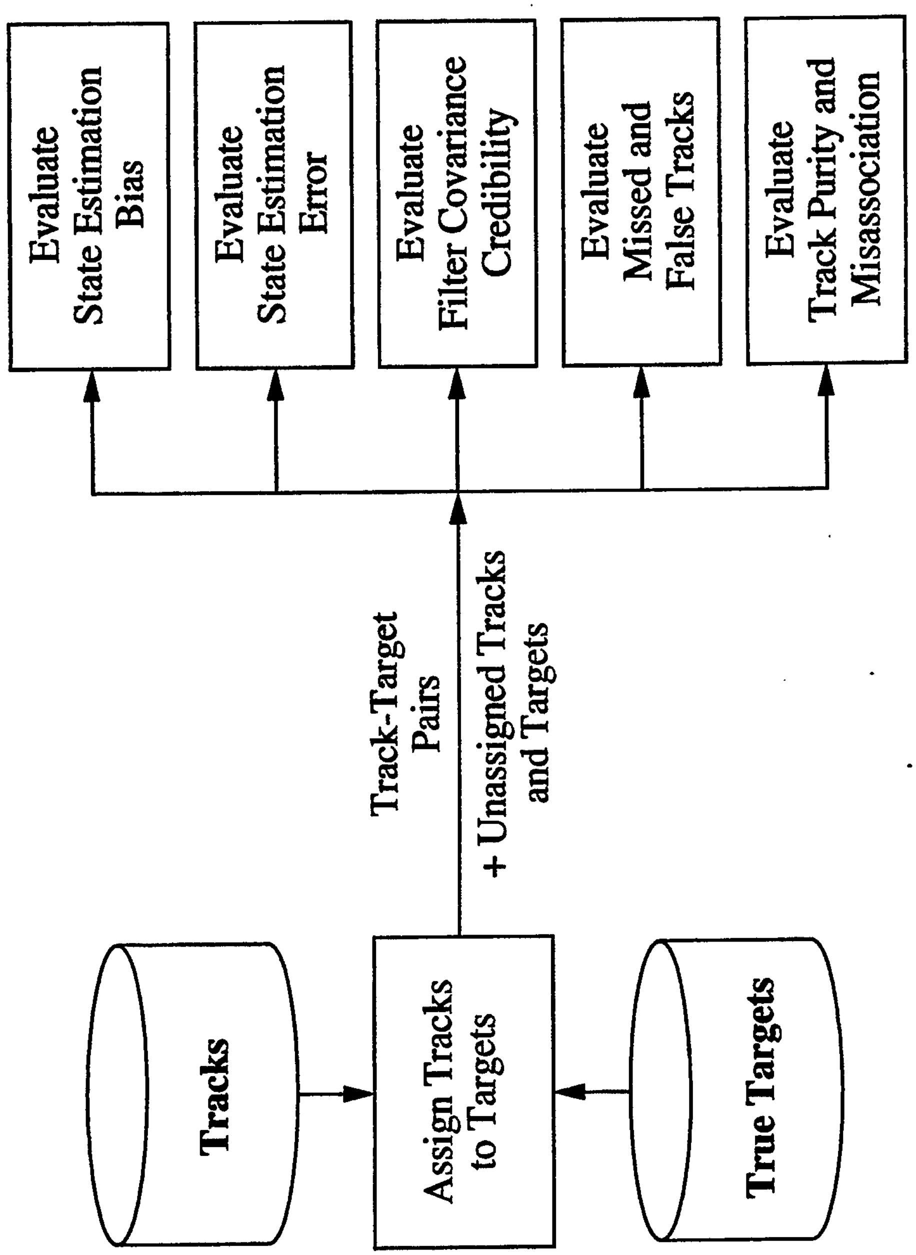

FIGURE 10 - A two-stage procedure to MSDF performance evaluation (Refs. 15 and 16)

A complete discussion of each of the two stages of the performance evaluation procedure described above (i.e., the track-to-target assignment process and the computation of measures of performance) will be presented in a separate document (Ref. 11).

4.6.3 Experimental Design (Evaluation Setup)

Aspects of "experimental design" include the formal methods of classical, statistical experimental design. Few if any MSDF research efforts use this type of strategy, presumably as a result of cost limitations. Nevertheless, there are the serious questions of sample size, confidence intervals for estimates, etc. to deal with in the formulation of any performance evaluation program since simple comparisons of mean values, etc. may not have very much statistical significance in comparison to the formal requirements of a rigorous experimental design (Ref. 13).

The combinations of candidates to be evaluated, scenarios to be applied and other independent variables define the test cases to be analyzed and are often expressed in a test matrix, as in Fig. 11. For each case ("bin") in the matrix, the set of independent variables ("factors") is defined and the model determines the measures of merit unique to that case. In the example shown in Fig. 11, eight sets of measures are computed for each candidate (Ref. 4).

4.7 Results Analysis

The analysis is performed by running the model for each case (full factorial experiment) or for carefully selected cases (partial factorial experiment) in the test matrix. The numerical results are then analyzed to determine the statistical significance of the results, often using the method of analysis of variance. In this method, a null hypothesis (which states that no difference exists between the populations of results for different treatments) is tested to determine the probability of rejection of the hypothesis as a function of the variances in the results of the experiments (e.g., analyses or simulation runs) (Ref. 4).

The analysis of data also includes the evaluation of results to:

- Define any uncertainties in the outcome.

- Specify the significance of marginal differences between performance in candidates as a function of input variables.
- Explain relations between output and input variables to understand the effects of complex processes within the model.
- Specify the sensitivity of measures of merit to input variables (scenarios, sensor suites, environment, algorithms, etc.).
- Recommend a candidate and provide the rationale for selection.

SCENARIO	1		2		3		4	
BENIGN / JAMMED MSDF SYSTEM CANDIDATE	B	J	B	J	B	J	B	J
A	1	2	3	4	5	6	7	8
B	9	10	11	12	13	14	15	16
C	17	18	19	20	21	22	23	24

TEST CONDITION 16 = SCENARIO 4
JAMMED ELECTRONIC ENVIRONMENT
MSDF SYSTEM CANDIDATE B

FIGURE 11 - Sample test matrix (Ref. 4)

5.0 CONCLUSION

DREV's Command and Control Information Systems Division is involved in multiple R&D activities in the field of local area Multi-Sensor Data Fusion (MSDF) for naval command and control afloat. The ultimate goal of these activities is to develop an optimal MSDF function that will enhance the tactical performance of the CCS of a CPF.

In this regard, the CASE_ATTI system is currently being used to support the development of advanced MSDF concepts that could apply to the current CPF sensor suite, as well as its anticipated upgrades, in order to improve its AWW performance against predicted future threats. This very practical study aims at identifying and developing techniques for combining radar/EO/ESM data, and at evaluating the real benefits of the combination. Two major aspects need to be addressed for this application. First, the representation of the actual CPF sensor suite to establish its baseline performance and, second, the quantification of the performance improvements gained when using an upgraded sensor suite combined with advanced MSDF concepts.

In this context, rigorously evaluating the performance of MSDF systems is of prime importance. An appropriate framework is required to guide the system designers in their evaluations of actual and future MSDF systems for the Canadian Forces. Unfortunately, no widely accepted scheme for characterizing the performance of MSDF systems is currently in use. Making a synthesis of the material currently available in the literature, this document proposed such a scheme (or methodology) to be used for MSDF performance evaluation in computer simulations.

In the perspective of evaluating their performance for surveillance and tracking, the use of sensors and advanced data fusion systems in naval defence applications was first discussed. The MSDF system was described in terms of its internal subprocesses, environment and output results in order to establish a solid basis for the rest of the document. Some of the performance evaluation issues encountered in the use of MSDF systems in defence applications were then addressed. In particular, some of the reasons why MSDF systems are hard to evaluate were discussed. The prediction of tracking performance using analytical evaluation methods was very briefly addressed. Discussions on the importance of Monte Carlo simulation techniques, and also on experimental trials for performance evaluation were then presented.

Evaluating MSDF systems is a complex and difficult activity. Haphazard approaches run the risk of introducing biases into the evaluation or generating completely spurious results. Rigorous and methodical testing is the only way to guarantee that the selected MSDF algorithms will work in a military application, and be the best choice from among several candidates. Therefore, a structured methodology was defined for evaluating and comparing alternative MSDF systems in a variety of common target environments. The proposed framework, described in the last chapter, comprises many facets. The performance evaluation activities are characterized by seven essential steps: 1) establishment of the context, 2) selection of a philosophy, 3) definition of specific objectives, 4) construction of alternatives, 5) selection of performance evaluation criteria, 6) development of an approach, and finally, 7) analysis of the results.

Because sensors and data fusion are contributors to improved information accuracy, timeliness, and content, a major objective of many fusion analyses is to determine the effect of these contributions to military effectiveness. However, the ability to relate MSDF performance to military effectiveness is difficult because of the many factors that relate improved information to improved combat effectiveness and the uncertainty in modeling them. A hierarchy of measures that relate performance characteristics of C3 systems (including MSDF) to military effectiveness was discussed. This hierarchy includes dimensional parameters, measures of performance, measures of effectiveness, and, for tactical MSDF systems, measures of force effectiveness.

A complete description of some important aspects such as the MSDF performance evaluation ambiguities, the two step procedure proposed to deal with these ambiguities (i.e., the track-to-target assignment process followed by the computation of measures of performance), implementation issues and details about each MSDF MOP, and the enhancement of CASE_ATTI to fully support advanced MSDF performance evaluation activities was beyond the scope of this document. Hence, the discussion of these issues has been deferred to a subsequent document (Ref. 11).

6.0 REFERENCES

[1] Roy, J.M.J., Bossé, É. and Des Groseilliers, L., "State of the Art in Local Area Sensor Data Fusion for Naval Command and Control Afloat," DREV R-9410, May 1995, UNCLASSIFIED

[2] Roy, J.M.J., Bossé, É. and Dion, D., "CASE_ATTI: An Algorithm-Level Testbed for Multi-Sensor Data Fusion", DREV R-9411, May 1995, UNCLASSIFIED

[3] Rothman, P.L. and Bier, S.G., "Evaluation of Sensor Management Systems", Proceedings of the IEEE 1989 National Aerospace and Electronics Conference NAECON 1989, Dayton, Ohio, May 22-26, 1989.

[4] Llinas, J. and Waltz, E.L., "Multi-Sensor Data Fusion," Artech House, Norwood, MA, 1990.

[5] Manson, K. and O'Kane, P.A., "Taxonomic Performance Evaluation for Multitarget Tracking Systems", IEEE Transactions on Aerospace and Electronic Systems, Vol. 28, No. 3, July 1992.

[6] Broida, T.J., "Performance Prediction for Multi-Sensor Tracking Systems: Kinematic Accuracy and Data Association Performance", Proceedings of the SPIE, Vol. 1198, "Sensor Fusion II: Human and Machine Strategies", 6-9 November, 1989.

[7] Wolff, T.O., Lutes, C.L., Mahler, R. and Fixsen, D., "Standards, Metrics, Benchmarks, and Monte Carlo: Evaluating Multi-Sensor Fusion Systems", Report, Unisys Defense Systems Inc., Eagan, Minnesota, 1991.

[8] Blackman, S.S., "Multiple-Target Tracking with Radar Applications," Artech House, Norwood, MA, 1986.

[9] Reid, D.B., "An Algorithm for Tracking Multiple Targets", IEEE Transactions on Automatic Control, Vol. AC-24, pp.843-854, December 1979.

[10] Fixsen, D., Lutes, C.L., and Mahler, R., "An Information Theory Based Figure of Merit For Target Tracking Algorithms", Report, Unisys Corporation, St-Paul, Minnesota, Revised Version: June 18, 1991.

[11] Roy, J.M.J. and Bossé, É., "A Two-Stage Procedure for Performance Evaluation of Sensor Data Fusion Systems in Computer Simulations", DREV Report to be published.

[12] Boily, E., "Performance Evaluation of a Multisensor Data Fusion Implementation: A Progress Report", SPIE Proceedings, Vol. 2235, "Signal and Data Processing of Small Targets 1994", 5-7 April 1994. (DREV Scientific Authority: J.M.J. Roy)

[13] Llinas, J., "Assessing the Performance of Multisensor Fusion Systems", SPIE, Vol. 1611, "Sensor Fusion IV", 1991.

[14] Hall, D.L., "Mathematical Techniques in Multisensor Data Fusion," Artech House, Norwood, MA, 1992.

[15] Drummond, O.E. and Fridling, B.E., "Ambiguities in Evaluating Performance of Multiple Target Tracking Algorithms", Proceedings of the SPIE, Vol. 1698, "Signal and Data Processing of Small Targets 1992", 1992.

[16] Fridling, B.E. and Drummond, O.E., "Performance Evaluation Methods for Multiple Target Tracking Algorithms", Proceedings of the SPIE, Vol. 1481, "Signal and Data Processing of Small Targets 1991", 1991.

[17] Drummond, O.E. and Blackman, S.S., "Challenges of Developing Algorithms for Multiple Sensor, Multiple Target Tracking", Proceedings of the SPIE, Vol. 1096, "Signal and Data Processing of Small Targets 1989", 1989.

[18] Drummond, O.E., "Multiple Target Tracking Lecture Notes", 1989.

[19] Chang, K.-C., Mori, S. and Chong, C.-Y., "Evaluating a Multiple-Hypothesis Multitarget Tracking Algorithm", IEEE Transactions on Aerospace and Electronic Systems, Vol. 30, No. 2, April 1994.

[20] Chang, K.-C., Mori, S. and Chong, C.-Y., "Performance Evaluation of Track Initiation in Dense Target Environments", IEEE Transactions on Aerospace and Electronic Systems, Vol. 30, No. 1, January 1994.

[21] Mori, S., Chang, K.-C., Chong, C.-Y. and Dunn, K.P. "Tracking Performance Evaluation: Prediction of Track Purity", Proceedings of the SPIE, Vol. 1096, "Signal and Data Processing of Small Targets 1989", 1989.

[22] Kovacich, M., "Track Acquisition Performance Analysis", Proceedings of the SPIE, Vol. 1954, "Signal and Data Processing of Small Targets 1993", 1993.

[23] Chang, K.-C., Mori, S. and Chong, C.-Y., "Performance Evaluation of a Multiple-Hypothesis Multi-Target Tracking Algorithm", Proceedings of the 29th IEEE Conference on Decision and Control, Honolulu, Hawaii, December 1990.

[24] Li, X. and Bar-Shalom, Y., "Stability Evaluation and Track Life of the PDAF for Tracking in Clutter", Proceedings of the 29th IEEE Conference on Decision and Control, Honolulu, Hawaii, December 1990.

[25] Bar-Shalom, Y., Editor, "Multitarget-Multisensor Tracking: Applications and Advances Volume II", Artech House, Norwood, MA, 1992.

[26] Paramax Systems Canada, "Implementation of Sensor-Level Hard Fusion for CSTSF Sensors: System Performance Specification", Preliminary, Defence Industrial Research Program, Agreement No.: 22072EV07, Financial Code: 1430AA-20JA-10226, Paramax DM No.: 112000012, June 11, 1993, UNCLASSIFIED. (DREV Scientific Authority: J.M.J. Roy)

[27] Sweet, R., "Command and Control Evaluation Workshop", MORS C2 MOE Workshop, Military Operations Research Society, January 1985.

INTERNAL DISTRIBUTION

DREV - R - 9423

1 - Director General
1 - Deputy Director General
1 - Director Command and Control Information Systems Division
1 - Director Electro-optics and Surveillance Division
6 - Document Library
1 - J.M.J. Roy (author)
1 - É. Bossé (author)
1 - G. Picard
1 - G. Otis
1 - J. Berger
1 - R. Carling
1 - B. Chalmers
1 - V. Pille
1 - D. Demers
1 - P. Labbé
1 - J.-C. Labbé
1 - D. Dion
1 - J. Boulter
1 - L. Sévigny
1 - J. Dumas

EXTERNAL DISTRIBUTION

DREV - R - 9423

2 - DSIS
1 - CRAD
1 - DGOR
1 - DMCS
1 - DMCS - 2
1 - DMCS - 4
1 - DMCS - 7
1 - DMFD
1 - DMOR
1 - DNR
1 - DNR - 4
1 - DNR - 8
1 - DRDCS
1 - DRDCS - 4
1 - DRDCS - 5
1 - DRDM
1 - DRDM - 6
1 - DRDM - 11
1 - DREO
Attn: D. Liang
1 - DREO
Attn: P. Yansouni
1 - DREO
Attn: R. Turner
1 - DREO
Attn: B. Wong
1 - Department of Electrical and Computer Engineering
Royal Military College of Canada
Kingston, Ontario
K7K 5L0
Attn: M. Farooq
1 - PMO Aurora

EXTERNAL DISTRIBUTION (contd)

DREV - R - 9423

1 - PMO Aurora
Attn: Lt. J.J.D. Lavoie

1 - PMO CPF

2 - CF Maritime Warfare School
CFB Halifax
Halifax, Nova Scotia
Attn: LCdr Connor
LCdr Harrison

1 - Paramax Systems Canada
6111, avenue Royalmount
Montréal, Québec
H4P 1K6
Attn: E. Shahbazian

1 - Thomson-CSF Systems Canada Inc.
18 Auriga Drive
Ottawa, Ontario
K2E 7T9
Attn: K. Bowring

1 - MacDonald Dettwiler
13800 Commerce Parkway
Richmond, British Columbia
V6V 2J3
Attn: B. Sharpe

1 - SPAR Aerospace Ltd.
9445 Airport Road
Brampton, Ontario
L6S 4J3
Attn: L.G. Gregoris

4 - Lockheed Canada Inc.
3001 Solandt Road
Kanata, Ontario
K2K 2M8
Attn: A. Sewards
W. Cornish
K. Doucet
R. Cooper

EXTERNAL DISTRIBUTION (contd)

DREV - R - 9423

1 - Software Kinetics Ltd.
65 Iber Road
Stittsville, Ontario
K2S 1E7
Attn: J. Bell

1 - Altek Systems Inc.
1643 Country Walk Drive
Orleans, Ontario
K1C 8E2
Attn: J.R. Altoft

2 - Department of Electrical Engineering
Laval University
Ste-Foy, Quebec
G1K 7P4
Attn: M. Lecours
D. Grenier

1 - Communication Research Laboratory
McMaster University
1280 Main Street West
Hamilton, Ontario
L8S 4L7
Attn: J. Litva

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.) DREV - 2459 Blvd Pie XI North, P.O. Box 8800, Courcelette, Qc, G0A 1R0	2. SECURITY CLASSIFICATION (overall security classification of the document. Including special warning terms if applicable) UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

"DEFINITION OF A PERFORMANCE EVALUATION METHODOLOGY FOR SENSOR DATA FUSION SYSTEMS" (U)

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

ROY, Jean, and BOSSÉ, Éloi

5. DATE OF PUBLICATION (month and year of publication of document) September 1995	6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.) 69	6b. NO. OF REFS (total cited in document) 27

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Report

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

N/A

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant) 0112C14A (Sensor Data Fusion)	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written) N/A
10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) DREV-R-9423	10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor) N/A

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

Unlimited Announcement

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

A key element in the anticipated information management problem on a naval platform is the ability to combine or fuse data, not only as a volume-reducing strategy, but also as a means to exploit the unique combinations of data that may be available. In this regard, the Command and Control Information Systems Division at the Defence Research Establishment Valcartier (DREV) is involved in multiple R&D activities in the field of local area Multi-Sensor Data Fusion (MSDF) for naval command and control afloat. The ultimate goal of these activities is to develop an optimal MSDF function that will enhance the tactical performance of the command and control system of a Canadian Patrol Frigate type of ship. In this context, rigorously evaluating the performance of MSDF systems is of prime importance. The system designer may have several MSDF concepts among which to choose in order to fulfill a particular requirement, the MSDF system may require optimization to improve performance, or the MSDF system may be undergoing testing to assure that it is operating correctly. Unfortunately, no widely accepted scheme for characterizing the performance of MSDF systems is currently in use. This document proposes such a scheme (or methodology) to be used for MSDF performance evaluation in computer simulations.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Performance Evaluation Methodology
Sensor Data Fusion

Requests for documents
should be sent to:

DIRECTOR SCIENTIFIC INFORMATION SERVICES
Dept. of National Defence
Ottawa, Ontario
K1A 0K2
Tel: (613) 995-2971
Fax: (613) 996-0392

#153574

NO. OF COPIES / NOMBRE DE COPIES	COPY NO. / COPIE N°	INFORMATION SCIENTIST'S INITIALS / INITIALES DE L'AGENT D'INFORMATION SCIENTIFIQUE
1	1	JC
AQUISITION ROUTE / FOURNI PAR ▶	DREV	
DATE ▶	18 Sep 95	
DSIS ACCESSION NO. / NUMÉRO DSIS ▶		

DND 1158 (6-87)

National Defence / Défense nationale

~~PLEASE RETURN THIS DOCUMENT TO THE FOLLOWING ADDRESS: DIRECTOR SCIENTIFIC INFORMATION SERVICES NATIONAL DEFENCE HEADQUARTERS OTTAWA, ONT. - CANADA K1A 0K2~~

~~PRIÈRE DE RETOURNER CE DOCUMENT À L'ADRESSE SUIVANTE: DIRECTEUR SERVICES D'INFORMATION SCIENTIFIQUES QUARTIER GÉNÉRAL DE LA DÉFENSE NATIONALE OTTAWA, ONT. - CANADA K1A 0K2~~

Toute demande de document
doit être adressée à:

DIRECTEUR - SERVICES D'INFORMATION SCIENTIFIQUE
Ministère de la Défense nationale
Ottawa, Ontario
K1A 0K2

Téléphone: (613) 995-2971
Télécopieur: (613) 996-0392