IDA

INSTITUTE FOR DEFENSE ANALYSES

Modeling the Observer in Target Acquisition

James D. Silk

April 1997

IDA Paper P-3102
Log: H 95-046893

This work was conducted under contract DASW01 94 C 0054, Assignment A-181, for the Defense Advanced Research Projects Agency. The publication of this IDA document does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official position of that Agency.

INSTITUTE FOR DEFENSE ANALYSES

IDA Paper P-3102

Modeling the Observer in Target Acquisition

James D. Silk

PREFACE

IDA prepared this paper in partial fulfillment of task order entitled "Target Acquisition and Search Studies," for Mr. Walter Hollis, Deputy Under Secretary of the Army [DUSA (OR)], and Mr. Raymond S. Balcerak, Program Manager, ARPA Infrared Focal Plane Array Producibility Microelectronics Technology Office.

IDA performed this work under the auspices of Mr. Hollis' ACQSIM (Acquisition Simulation Working Group) activity. In an earlier ACQSIM forum, Mr. David Dixon of TRAC-WSMR (TRADOC Missile Command–White Sands Missile Range) presented an analysis of alternative formulations of observer variability in target acquisition simulations. That analysis demonstrated the importance of resolving these long-standing issues, and motivated the present work.

The high quality target acquisition performance data which the Visionics Division of NVESD obtained in their Target Acquisition Model Improvement Program (TAMIP) Phase 1 target acquisition tests were indispensable to the completion of this study.

CONTENTS

FIGURES

TABLES

EXECUTIVE SUMMARY

The Night Vision Standard Model[1] defines P_∞ as the fraction of a large sample of observers, drawn from the standard ensemble, that will, given unlimited time, eventually detect a given target.

Even given that this definition is understood, what is the *meaning* of P_∞? The question has been asked frequently since the introduction of the Standard Model. At first glance, the question seems at best irrelevant; at worst, without content. After all, given the value of P_∞, we can by definition predict (up to statistical uncertainties) how many observers out of a given sample will eventually succeed in a specific target acquisition task.

Nevertheless, different interpretations are possible. A seminal work[2] on the subject discussed several of them at length and proposed possible implications. Here are two of these interpretations:

- All observers are equivalent in a statistical sense, but their responses to a given target are stochastic. Thus, if P_∞ is 50 percent, then any given observer has a 50-percent chance of detecting it. This is independent of that observer's chance of detecting any other $P_\infty = 50$ percent target.
- All observers are strictly ordered in target acquisition competence. There is no stochastic element at all; the detection "probability" is really just the fraction of the observer ensemble that is sufficiently competent to accomplish the detection of the given target. In this case, *every* $P_\infty = 50$ percent target is *always* detected by above-average observers and *never* detected by below-average ones.

It is well known that neither of these extreme interpretations is literally true. While some observers are certainly better than others, strict adherence to hierarchical ordering is not observed. So some sort of hybrid model seems to be called for. But to date no such approach has been validated, nor is there even an understanding of which approach is closer to reality.

1 Ratches et al.

2 Marta Kowalczyk and Stanley Rotman, *Extending the CNVEO Search Model to the Multitarget Environment*, Institute for Defense Analyses, IDA Paper P-2022, 1987.

It is easy to see why this matter has remained unresolved. For a one-on-one engagement the question of interpretation does not arise—both approaches give the same answer. Thus, there has in the past been only theoretical interest in determining the correct form of such a hybrid model. In fact, there has been little concern that the two most commonly used Army combat models, CASTFOREM and JANUS, adopt opposed views. CASTFOREM attributes the variability to the observers (observer-only draw), while the JANUS implicitly assumes that the target acquisition process is inherently stochastic (observer-target draw).

But, as recently demonstrated by David Dixon of TRADOC Analysis Command (TRAC), in many-on-many engagements, correlations between different observers does make a difference. Dixon's analysis gave new impetus to the resolution of the long standing question of the meaning of P_∞.

We believe that we have taken a large step toward the resolution of this problem via an analysis of NVESD observer test data. Our approach to the problem evolved considerably over time, as we learned which ideas were supported by the data and which were not. The final formulation of our model rests on the following assumptions:

- The probability that a given observer will detect a given target is a function of the sum of two variables. One of these variables is determined by the particular target signature, the other by the individual observer.
- The test data which we use to determine these parameters employed a representative sample of the standard observer ensemble.
- The estimate of the target signature variable that is computed in the wargames is not precisely the true value, but is an estimate subject to a quantitatively known modeling uncertainty.

The first two assumptions enable us to evaluate the relative importance of the stochastic and the observer components of the variability. Taken together, they support an important observation: the value of P_∞ is the average of the joint probability function over the observer distribution.

The third assumption points to a third source of variability in the simulation of target acquisition—namely, a modeling uncertainty—that we need to address. We have recently published a quantitative analysis of the uncertainty associated with model predictions.[3] This variability, it turns out, is at least as large as the other two. It is therefore important to simulate its effects.

Our final result is a complete model for the observer dependence of the target acquisition probability. All required functions and constants are provided; there are no free parameters. The explicit code we've included is suitable for guiding the model's implementation in the wargames. The approach is similar to the one now in use. In the setup phase of the simulation, random draws determine detection thresholds. But instead of a single random draw, there are three independent sets of draws:

- An observer-target draw, to account for the stochastic process;
- An observer draw, to account for the different level of capability of the observers; and
- A target draw, to account for the uncertainty in the target signature computation.

The essence of our work is the assumption that the target detection probability is determined by the sum of an observer variable and a target variable. This apparently innocuous assumption generates most of the simplifications that render the problem tractable. Therefore, it is this aspect of our approach that, we hope, will receive additional scrutiny. In particular, we look forward to alternative formulations which do not invoke this assumption.

Meanwhile, we offer our model as a practical approach that is sufficiently similar to existing approaches that it can be implemented relatively painlessly, while being as accurate as it is possible for such a simple model to be.

---

3 James D. Silk, *Statistical and Modeling Uncertainties in the Thermal TAMIP Predictions,* Institute for Defense Analyses, IDA Paper P-3078, 1995.

I. MOTIVATION

Stochastic wargames rely on randomness to rescue them from the plethora of unknowable variables that in reality affect the outcomes of hostile engagements. The intent is that by properly averaging over all possibilities, the most likely outcome can be determined. Unfortunately, the transition from a one-on-one model (even if it is well validated and understood) to one that accurately predicts many-on-many interactions is ambiguous.

Consider the following rather contrived example. Suppose that a red commando must avoid detection by both of two blue sentries to accomplish his mission. The one-on-one probability of detecting the commando (averaged over all observers) is assumed to be 50 percent. We further assume that the sentries are identically capable, co-located, and both fully vigilant. What is the commando's probability of success? Two methods of computation come to mind.

The assumption that is often invoked in elementary problems of this type is that the "and" condition—that is, the commando must evade both sentry one *and* sentry two—means we must multiply probabilities. The use of this recipe is based on the implicit assumption that the detections are statistically independent events, and are therefore uncorrelated. That is, the commando eludes the first sentry half the time, then he eludes the second sentry half of the remaining time, netting a 25-percent probability of success.

But the independence hypothesis may not be valid in fact. It is surely true that observers generally agree on the subjective evaluation of "easy" and "hard" targets. Indeed, it is possible to argue (in the absence of data to the contrary) that equally capable observers are always perfectly correlated in their responses. After all, the two sentries in our example receive identical stimuli; if the sentries are deterministic in their behavior, then perfect correlation of their responses is the logical conclusion. In this case, our commando succeeds 50 percent of the time, because if he eludes the first sentry then he is certain to elude the second.

(A third alternative, namely that the sentries are perfectly anticorrelated, is possible in principle. In this case, the commando never succeeds because if he gets by one, the other always detects him. It is inconceivable, however, that this is a property of real human sentries who receive identical stimuli.)

The two assumptions, uncorrelated observers and perfectly correlated observers, represent the extreme cases of a range of plausible models of human target acquisition that are partly deterministic and partly stochastic. It is interesting that the two most important Army force-on-force battlefield simulations, CASTFOREM and JANUS, use a different extreme case. In CASTFOREM, all of the variability is assumed to be due to variations among the observers, who are modeled as obeying a strict hierarchy; in JANUS, all observers are statistically equivalent, but independent. It is well known that neither is precisely correct, but up to now no method has been put forth for determining whether one is more correct than the other.

It might be hoped that the foregoing observations merely constitute a technical nuance that has no practical import. Recently, though, D. Dixon of TRAC/WSMR (TRADOC Analysis Command–White Sands Missile Range) clearly demonstrated that these two different assumptions can make a vast difference in the outcome of a simulated engagement. His examination of idealized forced march scenarios illustrated that the uncorrelated model (the so-called observer-target draw, which is used in JANUS) generates a kind of "stochastic fire control" relative to the perfectly correlated simulation (that is, a simulation using an observer-only draw, such as CASTFOREM). Thus, in the former case, one side would be completely attritted; in the latter, the fronts would march through each other.

As part of our Target Acquisition Model Improvement Program (TAMIP) and Acquisition Simulation Working Group (ACQSIM) activities, we have undertaken to resolve this issue. In Chapter II we examine the test data somewhat qualitatively to get an understanding of the nature of the correlations therein. In Chapter III we propose a model which incorporates an observer variable. Chapter IV presents the results of the data analysis. We pursue two versions of the analysis; the simpler approach is shown to give a preferable description of the data. Finally, in Chapter V we explicitly define a working algorithm for implementing a simulation of the proposed model.

II. OBSERVER TEST DATA

The data set we use herein is from the NVESD Phase 1a1 and 1b1 tests.[4] We have analyzed results from 17 observers on 313 targets in the former test, and 22 observers on 275 targets in the latter. For the sake of the discussions of this chapter only, we will examine a representative subset of the data comprising 8 observers and 10 targets from the 1a1 test.

For our purposes, the data consists of a matrix of zeros and ones. Each column of the matrix corresponds to a particular human observer who participated in the test as a subject. Each row corresponds to a particular target in one of a set of images that were presented to the subjects. The matrix entry in cell *i,j* is one if target *i* was detected by observer *j*, and zero otherwise. (The detection criterion used is 100 percent confidence that the object is a military target, with no requirement for any higher level of discrimination.) The matrix corresponding to our subset is:

Matrix A

1	1	1	0	0	0	0	0
1	0	1	1	1	0	1	0
1	1	1	0	0	0	0	0
1	1	1	0	1	0	1	1
1	1	1	0	0	0	0	0
1	1	1	1	0	1	0	0
0	0	0	1	1	0	1	0
1	1	0	0	0	1	0	0
1	1	1	0	1	1	1	1
1	1	1	1	0	1	0	0

If we average all the entries in the first column, we get an estimate of the probability that Observer #1 can detect a target from this set (90 percent). Similarly, averaging over the first row gives an estimate of the probability of detection of Target #1 (37.5 percent). It is hard to see any correlation in the above matrix. But if we reorder the rows and columns by descending total number of detections, we get Matrix B:

---

4 Barbara L. O'Kane, Clarence P. Walters, and John D'Agostino, "Report on Experiments in Support of Thermal TAMIP," NVESD, Fort Belvoir, VA, 1993.

Matrix B

1	1	1	0	1	1	1	1
1	1	1	0	1	0	1	1
1	1	1	1	0	1	0	0
1	1	1	1	0	1	0	0
1	0	1	1	1	0	1	0
1	1	0	0	0	1	0	0
1	1	1	0	0	0	0	0
1	1	1	0	0	0	0	0
1	1	1	0	0	0	0	0
0	0	0	1	1	0	1	0

This matrix looks as if it has some correlation among observers, but it is hard to quantify the degree of correlation. The correlation is obviously not perfect. If it were, then ordering the rows and columns would have eliminated all of the islands of zeros in ones, and vice-versa, as in the following fictitious data:

Matrix C

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0
1	1	1	1	1	1	0	0
1	1	1	1	1	1	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	1	0	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0

For perfect correlation, as in Matrix C, all of the following statements hold:

- It is possible to separate all the zeros from all the ones by drawing an ascending staircase through the data.
- If observer A is a better observer than B, then A can detect all the targets that B can, plus some others.
- Equivalently, if target X is easier than target Y, then X can be detected by all the observers that can detect Y, plus some others.

One can verify by inspection that these statements hold for the artificial data in Matrix C, but not for the real data in Matrix B.

Whichever phrasing one prefers, it is clear that the actual data does not conform to the perfectly correlated hypothesis. We now propose a means of quantitatively describing the intermediate case.

III. MODEL DEFINITION

A. CONCEPTUAL DESCRIPTION

The Thermal TAMIP target detection model predicts target detectability as a function of a composite statistic ("line pairs on target") that is computed from target properties. We refer to this and other models that do not explicitly include observer variables as target-only models. The model is stochastic in the sense that it provides outputs which are probabilities, and these probabilities are appreciably different from zero or one in a finite range of the target statistic. The actual relation between the P_D and the target statistic will be discussed later. For now, we simply characterize the function as in Figure III-1: P_D has a value of 0 for small values of the target statistic, 1 for large values, and an intermediate "stochastic" region in between, where it takes on values significantly different from 0 or 1.

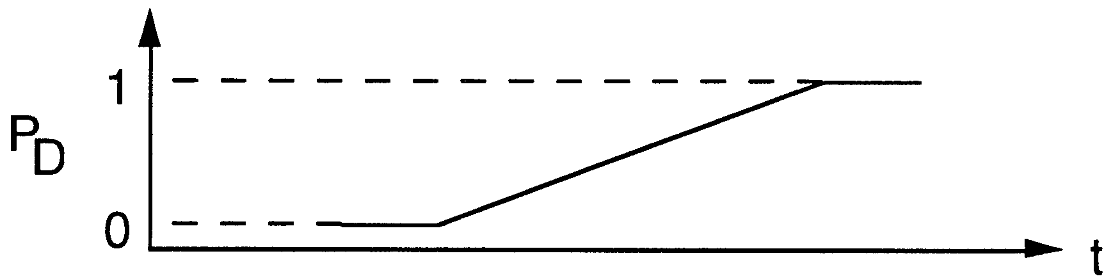


Figure III-1. Probability as a Function of a Target Statistic in a Target-Only Model

In the foregoing discussions, we have observed that there are two possible sources for this stochastic element. One source is the variability between human observers. Since individual observer properties are not specified, the output must refer to an average over a standard (but unspecified) observer ensemble. The other source is the target acquisition process itself. As we have seen, it seems to have some inherent element of randomness.

Our approach is to construct a model that nominally incorporates the observer variable. We use the term "nominally" because we do not seek to describe how it might be evaluated for a specific observer (e.g., a function of eyesight, experience, etc.). That is, we do not pretend to know how to compute the model inputs a priori. Instead, we simply use our construct to evaluate correlations between observers by taking an a posteriori approach: given the results of a test, what are the most likely values of the model inputs? What can be deduced from these values about the correlation between observers?

We assume that the proficiency of the subjects in a target acquisition task is uniquely determined by a single variable s. Similarly, we assume the detectability of the targets under a particular set of observational conditions is uniquely determined by the variable t. Then, a model estimate of the probability that a given target will be detected by a given observer must be some function, call it Π, of s and t. We assume that this functional relationship takes a certain specific but very plausible form, but for now, we discuss it in the same general terms that we used for P_D above.

We note that an elementary requirement for any such model is that we must by definition retrieve the target-only model when we average over the standard observer ensemble. In particular, this means that we shall need to find an explicit relationship between our target variable t and the target signature variable that appears in the Thermal TAMIP model.

To achieve these ends, we consider three models: the deterministic, the stochastic, and the hybrid. In the figures associated with each of these models, Figure III-1 is reproduced, to emphasize the requirement that we connect back to the "target-only" model that predicts performance.

1. Deterministic Limit

The first case that we consider is illustrated in Figure III-2. The two independent variables, s and t, are the vertical and horizontal axes, respectively. The shaded rectangle represents the ensemble of targets and observers that populate a hypothetical test. The probability function, Π, is defined on the plane. This dependent variable's axis is perpendicular to the page. The diagonal line divides the plane into two regions: to the upper right, where targets are easy and observers are good, $\Pi = 1$; to the lower left, $\Pi = 0$.

We call this version of the model "deterministic" because there is no region of the (s, t) plane where Π is different from 0 or 1. That is, for a given observer and a given target there is no random element. In this case the seemingly stochastic region of the "target-only" model arises entirely from the variation in the observer ensemble.

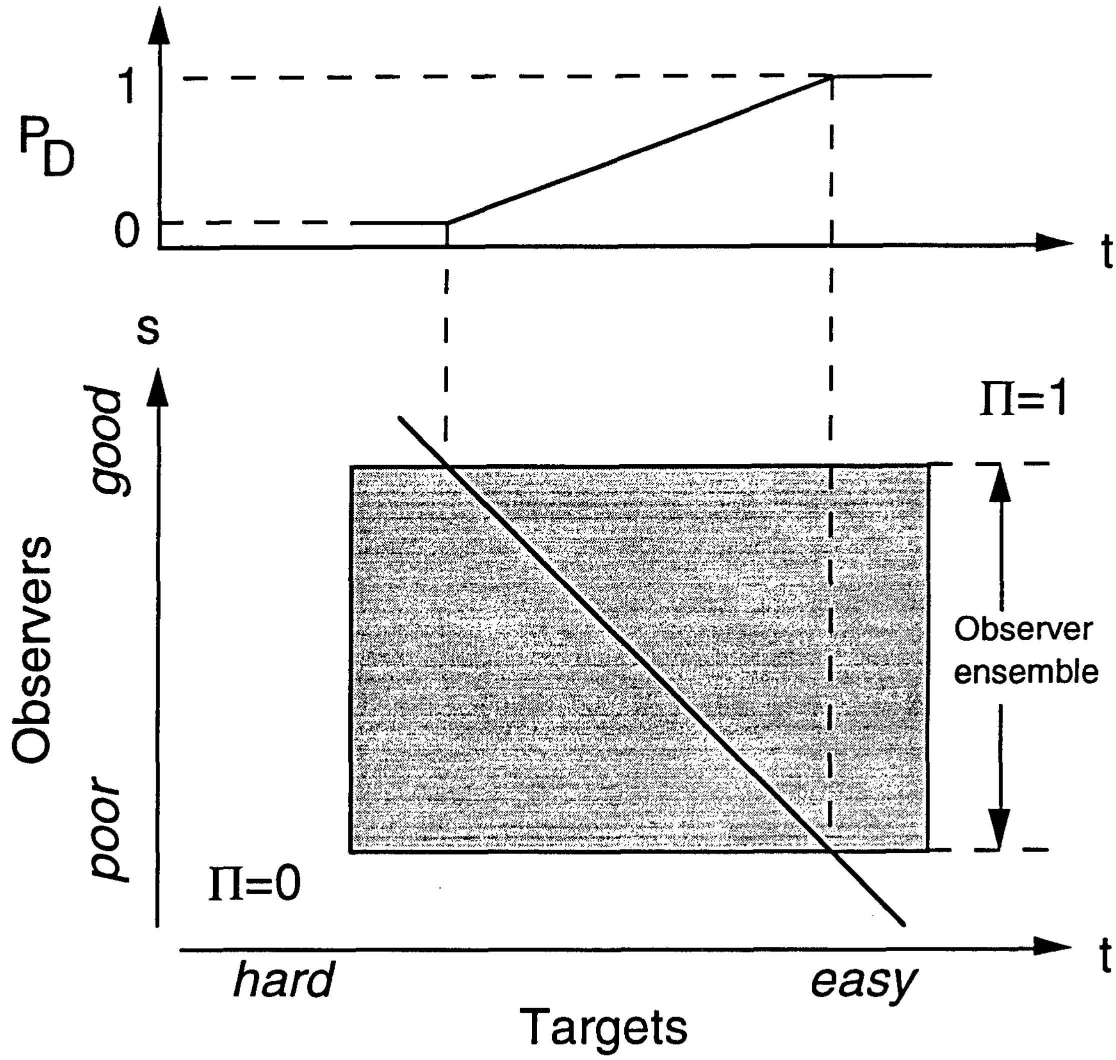


Figure III-2. Deterministic Model

2. Stochastic Limit

Figure III-3 depicts a different extreme. In this case all of the observers are essentially equivalent. This is depicted in the vanishing width of the shaded rectangle. On the other hand, the probability function Π now has an intermediate region—between the diagonal lines—where it takes on values significantly different from 0 or 1. All of the stochastic behavior of P_D is now attributed to the inherent randomness of the detection process itself. We therefore call this the stochastic limit.

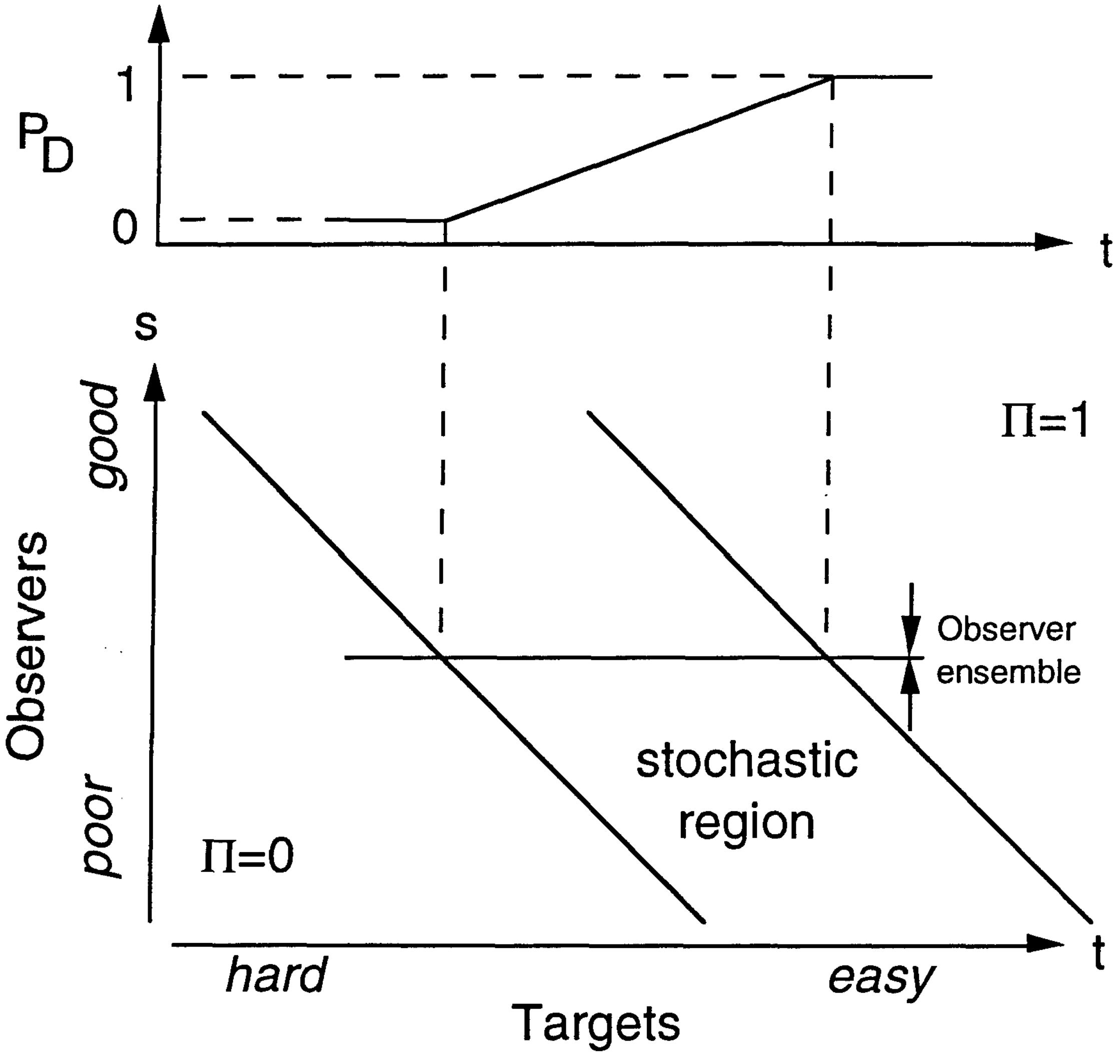


Figure III-3. Stochastic Model

3. The Intermediate Case

We have already observed that neither of these cases is correct. The deterministic version dictates a strict ordering of observer responses; we saw in Chapter II that this ordering is violated. The stochastic version is also wrong, since we know that there are significant variations within the observer population. Therefore, a hybrid of these alternatives is required.

Figure III-4 represents this hybrid model. As with the deterministic model there is variation within the observer ensemble. And there is a stochastic region for Π as well. The stochastic region of P_D reflects contributions from both effects. It is important to observe that the fixed width of the stochastic region of P_D places a constraint on the widths of the observer ensemble distribution and the stochastic region of Π, since these two functions must convolve to form P_D.

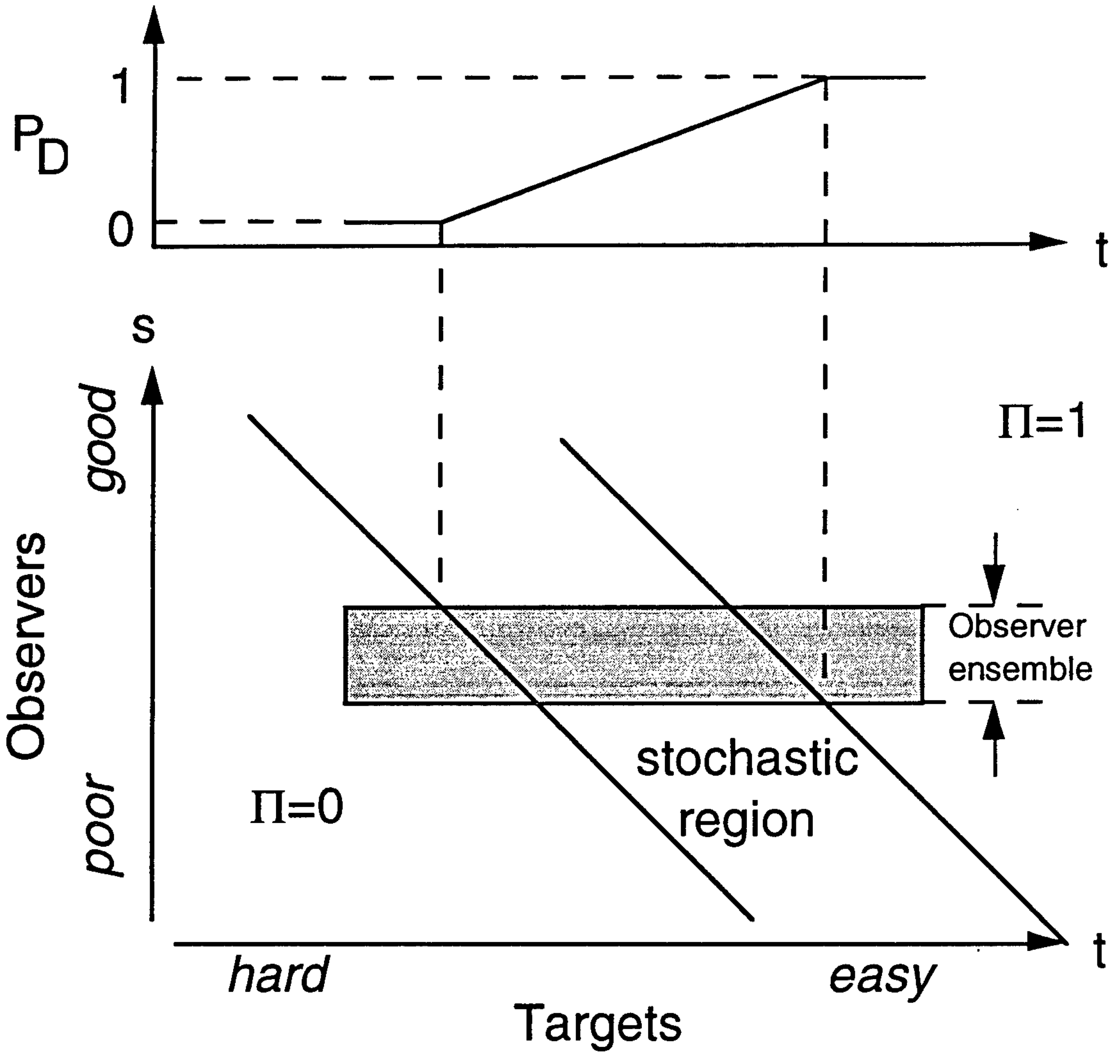


Figure III-4. Hybrid Model

B. FORMAL DEVELOPMENT OF THE MODEL

All of the conceptual ideas that we need have been discussed above. We now write the formulas explicitly so that we can pursue the data analysis required to validate the model and elucidate the model's implementation.

The form of the function that relates the target statistic to the probability of detection in the Thermal TAMIP Model is a logistic transform:

$$P_D(x) = L(b\,x) \quad ,$$

where

$$b = 3,\ x = \ln\left(\frac{M}{M_{50}}\right), \text{ and } L(u) \equiv \frac{1}{1+\exp(-u)} \quad ,$$

so

$$P_D = \frac{M^3}{M_{50}^3 + M^3} \quad .$$

Although the shape of this curve closely resembles the older version of the target transform probability function (TTPF), it has been chosen because it is better behaved analytically.[5] For the sake of consistency, we take our function Π to be a logistic function of the sum of the independent variables:

$$\Pi(s,t) = \Pi(s+t) = L(s+t) \quad .$$

In the data analysis that follows we shall derive values of s and t which optimize the fit to the field test data. First, let us consider the observer variable s.

It is not our intention to attempt to predict observer performance based upon a priori data such as eyesight, training, experience, etc. How, then, is the variable s useful? The variable s is a vehicle to connect to the standard observer ensemble. That is, there must exist—and we must specify!—some cumulative distribution function Ψ such that $\Psi(s)$ gives the fractile rating of the observer who is characterized by s. In short, the expression above for Π is of little use until we specify Ψ. Fortunately, in the design of the observer tests which were used for the development and validation of the Thermal TAMIP Model, great care was taken to use a representative sample of observer subjects. Therefore, we get the necessary information almost for free.

To get the functional form of Ψ we need to explicitly carry out the averaging over the observers. Let us define the function $\Phi(t)$ as the formal average of Π over the standard observer ensemble. (This new function is essentially P_D, but we have not yet connected t and x so we give it a unique name for now.) This average is simply a convolution integral of the observer-target function over the observer distribution:

$$\Phi(t) = \int \psi(s)\, \Pi(s+t)\, ds \quad ,$$

where ψ is the derivative of Ψ.

---

[5] In particular, its inverse can be written explicitly. It also has the useful property that it is related to its derivative by the formula $L' = L(1-L)$.

Let $\Phi(t) = L(at)$.[6] Again, this is consistent with the Thermal TAMIP definition. Then the functional form of ψ is determined by our choice of the logistic function for Φ and Π. Appendix A gives the method of derivation and illustrative computations. Remarkably enough, the result can be written in closed form:

$$\psi(s) = \frac{1}{2\pi a}\frac{\sin \pi a}{\cos \pi a + \cosh s}$$

$$\Psi(s) = \frac{1}{2} + \frac{1}{\pi a}\tan^{-1}\left(\tan\frac{\pi a}{2}\tanh\frac{s}{2}\right) \quad .$$

Note that Ψ, like L, is easily inverted. Thus, once we have determined a, we can readily construct a composite function that connects the observer fractile rank (R), the ensemble averaged target detectability (P_D), and the observer specific detectability (P):

$$P = \Pi(\Psi^{-1}(R), \Phi^{-1}(P_D)) \quad .$$

We now return to the target variable. To finally implement the model into the wargames, we need to make the connection to the "target-only" model (in our case, the Thermal TAMIP Model) by establishing the relationship between t and x. It seems reasonable at first to identify Φ with P_D and use the definitions above to establish the linear relationship between t and x. This yields

$$t = \frac{b}{a}x \quad ,$$

but there is a subtlety that we must address before we close out this discussion. A problem arises because the value of x computed in the Thermal TAMIP Model is actually an estimate of the true value of the target statistic—call it x'. We have shown elsewhere[7] that the data support the assumption that x' is an unbiased estimate of x such that $x = x' + \eta$, where η is a Gaussian random variable with zero mean and 0.36 standard deviation. See Figure III-5 for an illustration and further details. If this "ideal" value could be determined, the "exact" detection probability would be $P_D' = L(b'x')$ with $b' = 4.5$. Since t is a similarly "ideal" quantity that has been determined by inference rather than constructively, for the purposes of wargaming it is more appropriate to use

---

6 Since Π depends only on the sum of s and t, the zero of t was not fixed. By omitting a constant offset term in this equation, we are removing this ambiguity and constraining $t = 0$ for 50-percent targets. Equivalently, the 50th-percentile observer has $s = 0$.

7 See note 3.

$$t = \frac{b'}{a} x'$$

in the simulation.

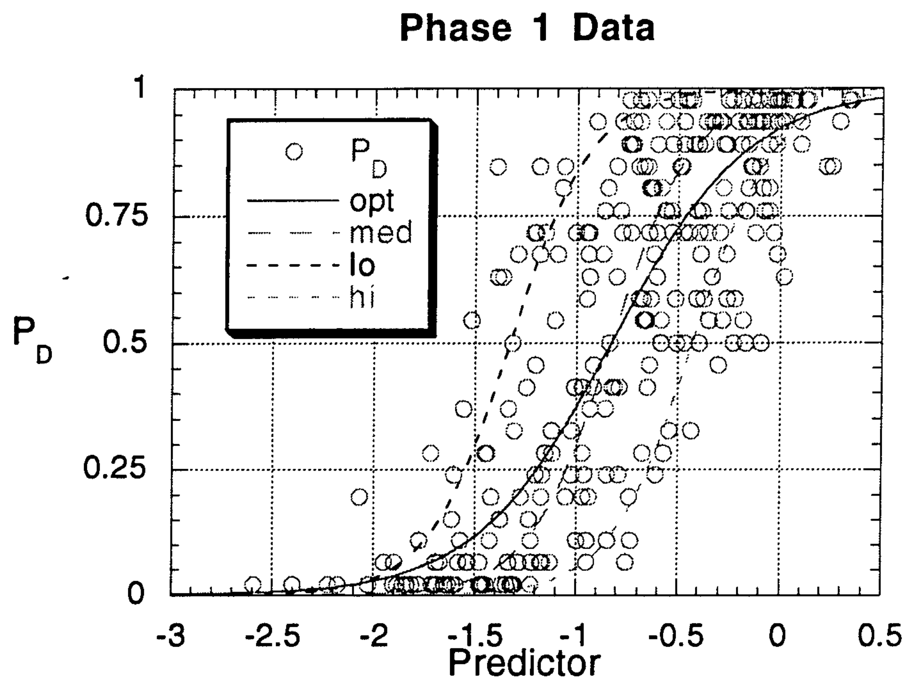


Figure III-5. Illustration of the Thermal TAMIP Fit to the NVESD Phase 1 Test Data. The shallow curve, labeled "opt," is the *b* = 3 fit, which is the result of minimizing the vertical errors. The curve labeled "med" is steeper, corresponding to *b'* = 4.5, and results from minimizing the horizontal errors. The other dashed curves illustrate the 80-percent confidence interval about "med."

C. COMMENTS ON THE DISTRIBUTION FUNCTIONS

It was necessary to be specific about the forms of the cumulative probability distribution functions (PDF's) that are derived above, since our objective is to implement a calculable model. [These PDF's are Ψ, Φ, and Π and their associated differential probability distribution functions (pdf's) ψ, ϕ, and π.] But in retrospect we suspect that all "reasonable" pdf's (i.e., bump in the middle, symmetric, roughly exponential tails) do about equally well. Our choices are a matter of taste and historical context rather than the result of an optimization analysis.

Similarly, the assertion that the target-only model is the average over the observer ensemble is almost without content. This is self-evident; there is nothing new here.

The important element of the model—the one that has predictive power and that we seek to test—is the conjecture that the observer-target model reduces to a function of a single variable, namely the sum of the observer and target statistics. This has an immediate and specific consequence: taking the derivative of the defining relationship above gives the relation among the pdf's:

$$\phi(t) = \int \psi(s)\,\pi(s+t)\,ds \quad .$$

It is now a simple matter to show that the model demands a constraint on the widths:

$$\sigma_{tgt}^2 = \sigma_{obs}^2 + \sigma_{obs \bullet tgt}^2 \quad ,$$

where $\sigma_{tgt} = \sqrt{\langle \Delta\phi^2 \rangle}$, etc. Note that the definition of P sets the scale of s and t; this fixes $\sigma_{obs \bullet tgt}$ at 1.814.

The data analysis will determine a value of t for each target and a value of s for each observer. Thus, the above constraint on the widths can be tested: the standard deviation of the observers' s values can be evaluated directly, yielding σ_{obs}. The value of a (which is the slope of the plot of t against the logodds of P_D) can be computed in a straightforward manner. Note that $1/a$ measures σ_{tgt} relative to $\sigma_{obs \bullet tgt}$. The value of $1/a$ will be unity in the stochastic limit and approach zero in the deterministic limit.

IV. DATA ANALYSIS

The objectives of the data analysis are:

- To determine model fits to the data and to show that the most general fitting procedure is extremely well approximated by a far simpler procedure;
- To show the consistency of the model across different observing conditions;
- To calibrate the model by evaluating the population statistics; and
- To validate the model by testing the constraint demonstrated at the end of the preceding chapter on the widths of the pdf's.

To meet these objectives we use the NVESD Phase 1 data set. For the purposes of this work, the NVESD Phase 1 observer tests each consisted of a set of observers attempting to detect a set of targets. Each observer was shown identical renderings of the same set of targets (but in different orders). We let the index i label the M targets, and j label the N observers. Then the results of the test are described by a rectangular matrix of M rows by N columns, called P. The elements of P are all zeros and ones (for misses and detects); a given element $P_{i,j}$ indicates the result of the jth subject's attempt to detect the ith target. We write $\Pi_{i,j} = \Pi(s_j,t_i)$ to denote that the matrix of model predictions for detection probabilities (on the left-hand side) can be computed from the model function if we know the $N + M$ values of s and t.

A. FITTING PROCEDURES AND RESULTS

For a given set of values for s and t, we can quantitatively compare the predictions to the data via the log-likelihood ratio:

$$LLR(s,t) = \sum_{i,j} \{P_{i,j} \ln(\Pi_{i,j}) + (1 - P_{i,j}) \ln(1 - \Pi_{i,j})\} \quad .$$

The best fit of the model to the data will correspond to the set of s_j and t_i for which this expression is maximized.[8]

[8] Harry F. Martz and Ray A. Walker, *Bayesian Reliability Analysis*, John Wiley & Sons, 1982.

We shall consider two approaches to this maximization procedure. First, we allow all of the s_j and t_i values to be independent and unconstrained. This means that for the NVESD Phase 1b data there are 275 + 22 = 297 free parameters. Second, we take a computationally less challenging approach by constraining the possible values of s_j and t_i in the manner described below. These constraints reduce the number of free parameters to two. This second approach is somewhat more instructive and yields greater insight into the model. Comparing the results of the two fitting procedures, it turns out that the simpler method costs almost nothing in terms of the quality of the fit. We believe that this approach to the data analysis will facilitate widespread confirmation of our results and will make the model easily supportable and applicable to different ensembles of observers.

1. Unconstrained Fit

The implementation of the first approach is straightforward in principle, but rather daunting in practice due to the high dimensionality. The usual method of choice (Mathcad™ for the Macintosh) was not practical at all. A means of searching on the s_j and t_i so as to maximize the *LLR* was developed (using Matlab™ on a Sparc10) and is given in Appendix B. The complete set was partitioned into four roughly equal pieces. Each of these pieces corresponds to a homogeneous clutter background, designated numerically from 1 to 4. Background 1 was the least challenging, 3 the most, with 2 and 4 the intermediate cases. We have broken the data out in this way in order to stress the model: it is conceivable that the best observers in one kind of clutter will not be the best in another, whereas the model would have the ranking of observers be independent (except for statistical variations) of the nature of the target acquisition task.

We have noted that Π is left unchanged if an arbitrary constant is added to all the *s*s and subtracted from all the *t*s. In the previous chapter we formally removed this ambiguity by insisting that the mean of s be zero. As a practical matter, it is easier to fix one observer to have $s = 0$ and leave the rest free; then implement the constraint after the fact.

2. Constrained Fit

Our second approach takes as its starting point the inverse of the equation that defines the logistic transform:

$$s + t = \ln \frac{\Pi}{1 - \Pi} \quad ,$$

where the function on the right is called the logodds of Π. In attempting to estimate the values of s and t that fit a given data matrix P, we might insert P on the right-hand side for Π and average over targets to get the values of s and over observers to get the values of t. The problem with this approach is that the matrix P consists entirely of 0s and 1s, so the logodds is undefined. Suppose instead that we effectively replace "the average of logodds" by the "logodds of the average." To do so, we employ the so-called uninformed estimate of the probabilities to preclude probability estimates that are zero or one. That is, we define

$$S_j = \frac{1}{M+1}\left\{\frac{1}{2} + \sum_{i=1}^{M} P_{i,j}\right\}; \quad T_i = \frac{1}{N+1}\left\{\frac{1}{2} + \sum_{j=1}^{N} P_{i,j}\right\} \quad .$$

With each of these probabilities we associate a logodds:

$$\sigma_j = \ln\frac{S_j}{1-S_j}; \quad \tau_i = \ln\frac{T_i}{1-T_i} \quad .$$

It is tempting to interpret these σs and τs as estimates of the ss and ts, which are the quantities that we seek. But this cannot be strictly correct, since these new quantities are strongly biased. That is, all of the observer odds could be driven to any arbitrarily large (or small) value by the inclusion of a suitable proportion of extremely easy (or difficult) targets.[9] By hypothesis, our ss (ts) are intrinsic properties of observers (targets) and should not be affected by the composition of a particular test.

But this is exactly the price we expect to pay for replacing the average of a function by the function of an average. We know that there is a relationship between the two, but the relation depends on the underlying populations—which for targets is scenario dependent, and for observers (at least for now) is not specified. As a way to undo this damage, we presume that the ss and ts are linear transforms of the experimental logodds (i.e., the σs and τs), and that the specific relationship is dependent on the experimental composition. To be precise, we assume that the unbiased variables can be written in terms of the biased ones,

$$s_j = \beta_1\left(\sigma_j - \langle\sigma\rangle\right), \; t_i = \beta_2 \tau_i \quad ,$$

with the understanding that the validity of our conjecture is to be established in the data analysis which follows. As we noted previously, the ss and ts are only determined up to a

9 On the other hand, recall that for tests of interest to us, the observer sample is representative of the standard ensemble. Therefore, the t values are, by hypothesis, independent of the particular test.

relative additive constant. In the above expression we implicitly define that constant such that our average observer will have $s = 0$, and 50-percent targets will have $t = 0$. Also note that the parameter β_2, which is determined in the fit, is exactly the value $1/a$, the importance of which is discussed at the end of the preceding chapter.

We have thus reduced the number of free parameters needed to fit the NVESD Phase 1 data set from hundreds to two. As a means of fitting data, this choice is far simpler than the more general method and is to be preferred, provided that the quality of the fit is not much worse. As we shall see, there is almost no penalty for this simplification.

The Mathcad™ script used to perform these computations is shown in Appendix C.

3. Comparison of Results

The quality of the fits is determined by the log-likelihood ratio (LLR). Table IV-1 is a comparison of the two methods. A general rule of thumb is that an additional degree of freedom (DF) yields significant improvement if it improves the LLR by at least one unit. We see that in the best case, the 294 additional free parameters yields an improvement of 2.49 units. Another way to look at this is that the specification of these 294 extra parameters has yielded an additional 2.49/ln(2) = 3.6 bits of useful information. Clearly, the extra effort is not worth the trouble.

Table IV-1. Log-Likelihood Ratio Comparison Between Two Model Implementations by Clutter Class

Clutter Class	LLR		
	Method 1 (296 DF)	Method 2 (2 DF)	Difference
1	− 402.44	− 403.40	0.96
2	− 377.93	− 380.42	2.49
3	− 454.50	− 456.88	2.38
4	− 473.97	− 475.87	1.90

The comparison of the fitted s values further illustrates the near equivalence of the two sets of results. Figure IV-1 shows the results for Clutter Classes 2 and 3. According to the table, these cases should show the most dissimilarity between the two methods. In fact, there is very little difference between the two methods, with most of the differences being associated with the outliers.

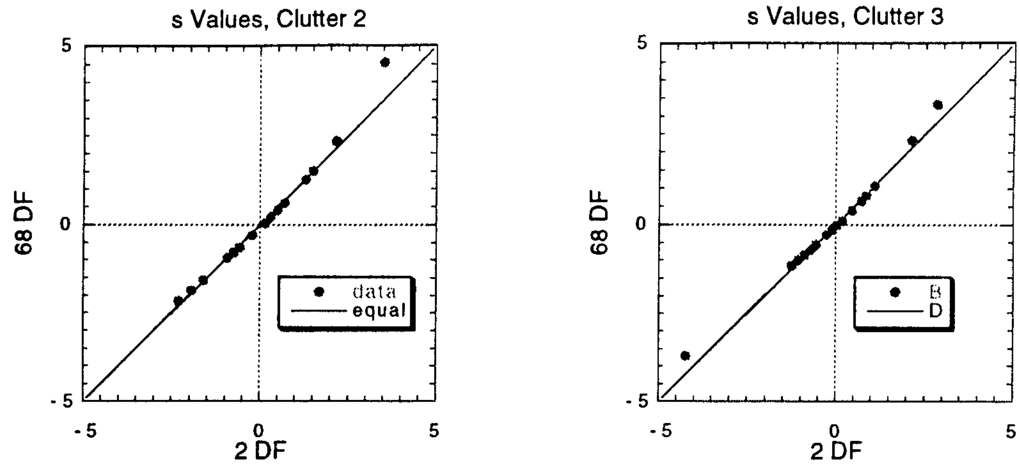


Figure IV-1. Equivalence of Observer Fits for Vastly Differing Degrees of Freedom

B. MODEL CONSISTENCY

Our model assumes that the competence of observers, which is measured by the variable *s*, is independent of the specific target acquisition task that is attempted. Thus, if our model is correct, we expect that it will produce roughly the same *s* value for each observer, even if we change the level of difficulty of the test.

In Figure IV-2 we compare the values of *s* that were determined for the four clutter classes in the test. (For reference, the average P_Ds for each clutter class are given in Table IV-2 below.) The figure shows two representative comparisons of the six possible. It is clear that the data follow the suggested trend.

However, while we have not performed a full test of significance, we believe that there is sufficient departure from out prediction to indicate that out hypothesis is not 100-percent correct. This is not surprising; we fully expect that military observers are not as one-dimensional as this model pretends. Nevertheless, the degree of correlation in the figure is encouraging. We will show that our approximation is good enough for our purposes, even though it does not hold precisely.

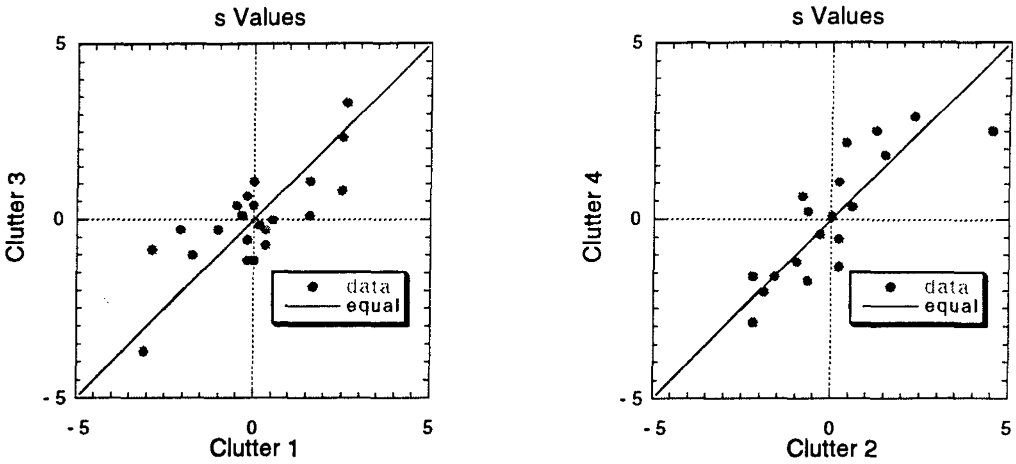


Figure IV-2. Correlation of Observer Performance in Various Clutter Conditions

C. CALIBRATION OF MODEL STATISTICS

The important parameters that need to be documented are the widths σ_{tgt}, σ_{obs}, and $\sigma_{obs \cdot tgt}$ of the functions ψ, ϕ, and π. Recall that by definition we have set $\sigma_{obs \cdot tgt}$ to 1.814, so we need to determine the other two from the data.

The value of σ_{tgt} is determined by the relation $\sigma_{tgt} = \beta_2 \sigma_{obs \cdot tgt}$, which holds as a direct consequence of the two-parameter fitting methodology. The value of σ_{obs} is simply the standard deviation of the fitted *s* values. This follows from our assumption that the observers in the test constitute a representative sample of the standard ensemble.

Table IV-2 shows that there is good consistency among the standard deviations for the four clutter classes. The global fit, to all data simultaneously, gives somewhat smaller width values, consistent with our observation that there is slight reordering of observers as the difficulty varies. When all levels of difficulty are folded together, the observer ensemble looks slightly more homogeneous.

Table IV-2. Net Detection Probabilities and Standard Deviations by Clutter Class

Clutter Class	Net P_D	σ_{obs}	σ_{tgt}
1	.77	1.49 ± .11	2.46 ± .13
2	.61	1.38 ± .12	2.38 ± .14
3	.35	1.37 ± .12	2.33 ± .14
4	.44	1.57 ± .11	2.56 ± .15
global fit	.54	1.27 ± .05	2.32 ± .07

D. VALIDATION OF MODEL CONSTRAINTS

The essential test of the model is the consistency between the widths of the *s* and *t* distributions that was shown previously: $\sigma^2_{tgt} = \sigma^2_{obs} + \sigma^2_{obs \bullet tgt}$. This goes to the central assumption of the model—that target detection depends on *s* + *t* only.

Figure IV-3 shows the result of this test. The points represent the data of Table IV-2. They are individually consistent with each other (marginally), and with the curve that is generated by the constraining relation above. Collectively, the data points all lie above the curve, which is probably an indication that our central assumption is violated at some level of precision. As noted above, this should not surprise us. Rather, it is remarkable that the model works as well as it does, given the simplicity of the assumptions.

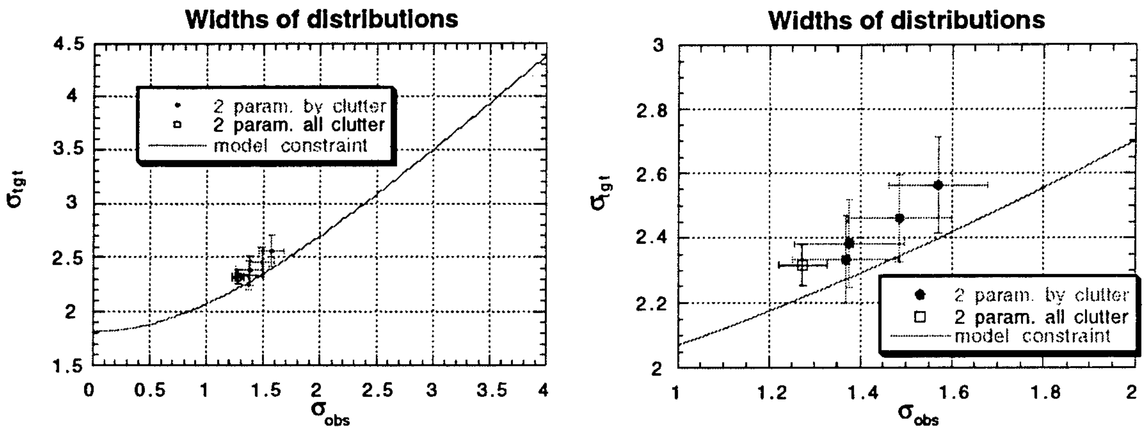


Figure IV-3. Consistency of Results for Different Clutter Conditions and Adherence to Model Constraint

The global fit indicates that the ratio of the widths $\sigma_{obs}:\sigma_{obs \bullet tgt}:\sigma_{tgt}$ is close to that of a 3:4:5 right triangle. This is the ratio we shall take as our baseline.

The knowledge of these parameters enables us to address questions about individual observer performance. Figure IV-4, for example, shows the expected performance of individual observers, given targets whose ensemble averaged P_D is known. For example, for a P_D = 50-percent target, the 90th-percentile observer detects with 75-percent probability while the 10th-percentile observer has only a 25-percent chance.

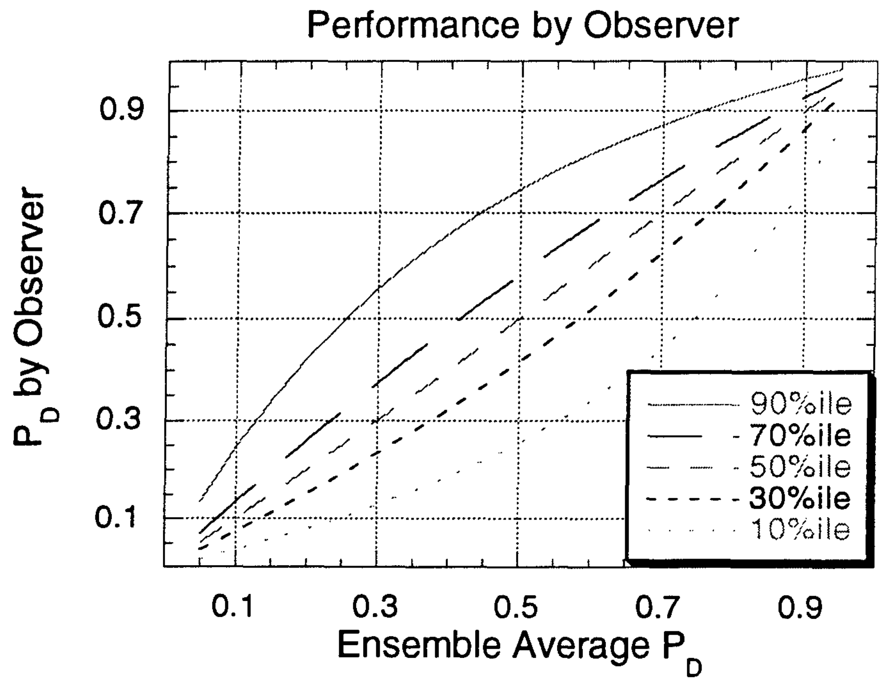


Figure IV-4. Performance of Individual Observers by Rank

V. SIMULATION

The foregoing methods and results are summarized in this chapter. We do so by putting them together in a working simulation that properly accounts for the observer variability and the inherently stochastic component of target acquisition, as well as the modeling uncertainty. The latter is a somewhat separate issue, treated here mainly for the sake of completeness and consistency.

We assume that the wargame contains a target-only acquisition module. By this we mean a model like Thermal TAMIP that, for given target signature, outputs the ensemble average probability that the target will be detected. Having validated herein that the target-only probability is a convolution of the observer-target kernel with the observer ensemble distribution, there are only three independent numbers that are needed for the simulation:

- The width of the target-only probability function in units of the "observer-target" function. This is the value $1/a$. Our baseline ratio from the data analysis is $\sigma_{obs}:\sigma_{obs \cdot tgt}:\sigma_{tgt} = 3:4:5$, so $1/a=5/4$.
- The scale factor that connects our t variable to the Thermal TAMIP x variable. We initially set the width of the observer-target function to 1.814, an arbitrary choice. The relative scale is defined by the relationship $at = Ex$, where E is the Thermal TAMIP exponent; we have shown[10] that the scale factor has the value $E / a = 5.625$.
- The standard deviation of the random variable associated with the uncertainty in the target-only model. In the x scale, this value is 0.38.

In implementing the simulation we take the same strategy that is used in JANUS and CASTFOREM. The random draws are done initially and are used to construct detection thresholds. Then in the course of the wargame, detection is declared (or is not) based upon whether the model produced probability exceeds (or does not) the predetermined threshold. The proper implementation of our modeling efforts is contingent upon correct construction of the thresholds.

---

10 See note 3.

We find that we need three independent draws:

1. An observer draw, to reflect the observer rank;
2. An observer-target draw, to represent the stochastic component of the acquisition process; and
3. A target draw, to simulate the uncertainty in the target-only model prediction.

The distribution functions are completely determined by the three values that we have in hand. They are all invertible in terms of elementary functions (we assume a logistic distribution for the modeling uncertainty), so the thresholds are easily obtained. Once the thresholds are computed, the game is played precisely as it was previously.

The Mathcad™ script that follows illustrates the simulation.

File sim.mcd
Simulates the observer target model.

Equations with asterisks * are not used, but are for reference or clarity.

Preliminary function definitions:

$\text{rnrnd}(m, s) := s \cdot \sqrt{-2 \cdot \ln(\text{rnd}(1))} \cdot \cos(\text{rnd}(2 \cdot \pi)) + m$ * — Gaussian random numbers.

$\text{logit}(x) := (1 + \exp(-x))^{-1}$ * — Logistic transform.

$\text{logodds}(P) := \ln\left(\frac{P}{1-P}\right)$ — Inverse of logistic transform.

We take the relative widths of the distribution functions from the data analysis of Chapter 4:

$a := \frac{4}{5}$ — Assume baseline 3:4:5 ratio. See text.

$r := \frac{a}{1}$ — Notational convenience.

$w_L := \sqrt{\int_{-20}^{20} \frac{x^2 \cdot \exp(x)}{(1 + \exp(x))^2}\, dx}$ — Width of the Logistic pdf.

$\sigma_{ObsTgt} := w_L$ — By definition; this defines our scale. Below we connect to the usual NVL Model variables.

$\sigma_{Tgt} := \frac{1}{a} \cdot \sigma_{ObsTgt}$

$\sigma_{Obs} := \sqrt{\sigma_{Tgt}^2 - \sigma_{ObsTgt}^2}$

Now need to relate the s, t variables to the scale of x and consider uncertainties in the target-only model:

$M_{50} := 1$ *

Option 1: Assume modeling is precise:

$E := 3$

$x(M) := \ln\left(\frac{M}{M_{50}}\right)$ *

Option 2: Assume modeling uncertainty is as given in IDA Paper P-3078:

$E' := 4.5$

$\sigma_x := 0.38$

$x'(M) := x(M) + \text{rnrnd}(0, \sigma_\eta)$ *

The scale factor relates the two sets of variables; the defining relation is a*t=E*x or =E'*x'. We choose Option 2.

$P_D(M) := \text{logit}(E' \cdot x'(M))$ *

$\text{scale} := \frac{E'}{a}$

$\sigma_{Unc} := \text{scale} \cdot \sigma_x$

Define probability functions:

$\text{Inv\_Stochastic\_CDF}(P) := \text{logodds}(P)$

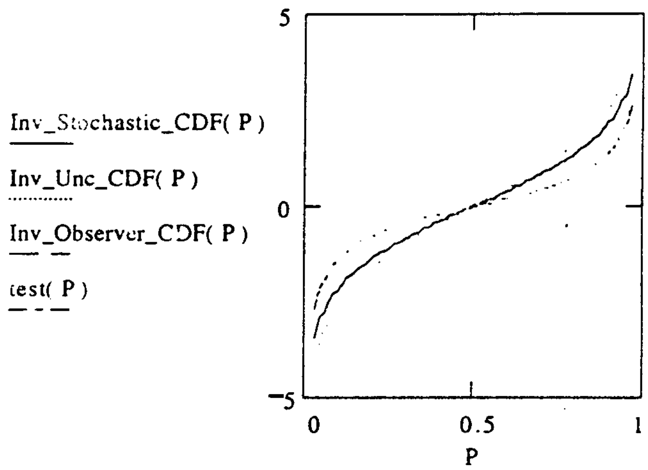
$\text{Inv\_Unc\_CDF}(P) := \dfrac{\sigma_{Unc}}{w_L} \cdot \text{logodds}(P)$ Will use this instead of inverse error function.

$\text{Inv\_Observer\_CDF}(P) := \dfrac{2}{a} \cdot \text{atanh}\left[\dfrac{\tan(\pi \cdot r \cdot (P - 0.5))}{\tan(\pi \cdot r \cdot 0.5)}\right]$

$\text{test}(P) := -\dfrac{1}{a} \cdot \ln\left(\dfrac{\sin(\pi \cdot r)}{\tan(P \cdot \pi \cdot r)} - \cos(\pi \cdot r)\right)$ Alternate form for inverse observer CDF.

Now we can look at the distribution functions and their width parameters.

$P := 0.03, 0.04 .. 0.97$



$\sigma_{ObsTgt} = 1.814$

$\sigma_{Tgt} = 2.267$

$\sigma_{Obs} = 1.36$

$\sigma_{Unc} = 2.138$

Following the current practice in CASTFOREM and JANUS, we run the simulation in two steps. First, before the engagement begins we roll dice to determine detection thresholds. Then we begin the game, declaring detections when the detection probability from the target-only model exceeds the preset threshold.

First do the preliminary draws. Need obs, tgt, and obs•tgt draws...

$n\_obs := 27$ $n\_tgt := 25$

$i\_obs := 0..n\_obs - 1$ $i\_tgt := 0..n\_tgt - 1$

$Observer\_Fractile_{i\_obs} := rnd(1)$ This determined the observer's fractile rating.

$Target\_Fractile_{i\_tgt} := rnd(1)$ This is the fractile rank of the target-only model uncertainty.

$Stochastic\_Fractile_{i\_obs,\, i\_tgt} := rnd(1)$ And this is the inherently stochastic piece of the target acquisition process.

Do the inverse transforms.

These manipulations just extend the rows/cols so all matrices are the same dimension:

$s1\_thresh := \overrightarrow{Inv\_Observer\_CDF(Observer\_Fractile)}$ $c1_{i\_tgt} := 1$ $s\_thresh := s1\_thresh \cdot c1^T$

$t1\_unc := \overrightarrow{Inv\_Unc\_CDF(Target\_Fractile)}$ $c2_{i\_obs} := 1$ $t\_unc := c2 \cdot t1\_unc^T$

$st\_thresh := \overrightarrow{Inv\_Stochastic\_CDF(Stochastic\_Fractile)}$

Verify the statistics generated by the random draws. Note the sample is small...

$\sigma_{Obs} = 1.36$ $stdev(s1\_thresh) = 1.351$

$\sigma_{Unc} = 2.138$ $stdev(t1\_unc) = 2.022$

$\sigma_{ObsTgt} = 1.814$ $\frac{1}{n\_tgt} \cdot \sum_{i\_tgt} stdev\left(st\_thresh^{\langle i\_tgt \rangle}\right) = 1.767$

Construct final threshold values.

$t\_thresh := st\_thresh - s\_thresh - t\_unc$

$P\_thresh := \overrightarrow{logit(a \cdot t\_thresh)}$ * Just a reminder.

$x\_thresh := \frac{1}{scale} \cdot t\_thresh$ Transform to x-space for direct comparison to Thermal TAMIP output.

Now play the game. We simulate acquisition for all observer-target pairs:

$$R_{i\_obs,\, i\_tgt} := \sqrt{0.5^2 + 5^2 \cdot \left(\frac{i\_obs}{n\_obs} - \frac{i\_tgt}{n\_tgt}\right)^2}$$

Two lines 5 km long, 0.5 km apart.

$$x := \overrightarrow{\ln\left(\frac{1}{R}\right)}$$

Signatures otherwise identical.

$$D := \overrightarrow{(x > x\_thresh)}$$

Compare to threshold values.

The results of the engagement:

$$D = \begin{bmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1
\end{bmatrix}$$

GLOSSARY

ACQSIM	Acquisition Simulation Working Group
ARPA	Advanced Research Projects Agency
DF	degree of freedom
LLR	log-likelihood ratio
NVESD	Night Vision and Electronics Systems Directorate
pdf	probability distribution function
TAMIP	Target Acquisition Model Improvement Program
TRAC-WSMR	TRADOC Analysis Command – White Sands Missile Range
TTPF	target transfer probability function

APPENDIX A

DERIVATION OF OBSERVER FUNCTION

APPENDIX A
DERIVATION OF OBSERVER FUNCTION

The form of the probability distribution function (pdf) that describes the observer ensemble is uniquely determined by our choice of the logistic function for the target-only and observer-target probability functions. In this appendix we explicitly derive this form of the observer ensemble distribution function.

We emphasize that the precise forms of the functions that we introduce are not of fundamental importance. Since the current state of the art does not support excursions into the tails of the probability distributions, one "S-shaped curve" is more or less as good as another, as long as the important parameters (namely the position and width of the transition region) are accurately represented. We have been explicit because it is necessary to generate a concrete model which is suitable for implementation into the wargames and at the same time is consistent with the existing suite of model results. Additional work on the specific functional forms may lead to marginally better fits, but will likely be both intellectually and practically unproductive.

The logistic function that defines the observer-target probability function is

$$F_a(x) \equiv \lambda(ax) = \frac{1}{2}\left(1 + \tanh\frac{ax}{2}\right) \quad ,$$

and a similar definition holds for the target-only function, $F_b(x)$. The defining relation is a convolution integral; our problem is to find $K_{a,b}(x)$ such that

$$F_b(x) = \int_{-\infty}^{\infty} K_{a,b}(x - x')F_a(x')dx' \quad .$$

It will be convenient to use Fourier transforms, so we need to work with normed functions. Taking the derivative yields

$$f_b(x) = \int_{-\infty}^{\infty} K_{a,b}(x - x')f_a(x')dx' \quad ,$$

where

$$f_a(x) \equiv \frac{d}{dx}F_a(x) = \frac{a}{4}\operatorname{sech}^2\frac{ax}{2}$$

and f is normed to unity.

Taking the Fourier transform and solving for the transformed kernel gives

$$\hat{K}_{a,b}(\xi) = \hat{f}_b(\xi)/\hat{f}_a(\xi),$$

where the hat denotes the Fourier transform, ant it can be shown that[1]

$$\hat{f}_a(\xi) = 2\frac{\pi\xi}{a}\operatorname{csch}\frac{\pi\xi}{a} \quad .$$

Then

$$\hat{K}_{a,b}(\xi) = \frac{a \sinh \frac{\pi\xi}{a}}{b \sinh \frac{\pi\xi}{b}}$$

and the inverse transform yields[2]

$$K_{a,b}(x) = \frac{b}{2\pi r}\frac{\sin \pi r}{\cos \pi r + \cosh bx} \quad ,$$

where $r = b/a$.

The cumulative distribution function (cdf) can also be written in terms of elementary functions[3]:

$$H_{a,b}(x) \equiv \int_{-\infty}^{x} K_{a,b}(x')dx' = \frac{1}{2} + \frac{1}{\pi r}\operatorname{atan}\left(\tan\frac{\pi r}{2}\tanh\frac{bx}{2}\right) \quad .$$

Note that the cdf depends only on the ratio $r = b/a$, as we would expect. The final result for the observer ensemble cdf is then

$$\Psi(s) = H_{1,b}(s) \quad ,$$

where $b = \frac{1.814}{\sigma_s}$

and σ_s is the standard deviation of the observer *s* values, as determined in the experimental data analysis.

---

1 A. Erdelyi et al., "Tables of Integral Transforms," Volume 1, McGraw-Hill, 1954, p. 30, equation (2).

2 Ibid., equation (6).

3 I.S. Gradshteyn and I. M. Ryzhic, "Tables of Integrals, Series, and Products," Academic Press, 1965, p. 108, equation 2.444.1.

The following two Mathcad™ scripts verify these results numerically. The file conv.mcd explicitly carries out the convolution and compared to the expected result, then explores the shape of the function and verifies the quadratic relation among the width parameters. The file cdf.mcd confirms that H is indeed the cdf associated with K.

conv.mcd

Verify that a kernel of the form K convolved with f results in a rescaling of f.
from Bateman's Tables of Integral Transforms, A. Erdelyi, page 30 equations (2) and (6).

$$f(a,x) := \frac{a}{4}\cdot \operatorname{sech}\left(a\cdot\frac{x}{2}\right)^2$$

Define f. Note a is inverse length scale.

$$K(a,b,x) := \frac{a}{2\cdot\pi}\cdot\frac{\sin\left(\pi\cdot\frac{b}{a}\right)}{\cosh(b\cdot x) + \cos\left(\pi\cdot\frac{b}{a}\right)}$$

Define K.
Must have b<a because output f is wider after convolution.

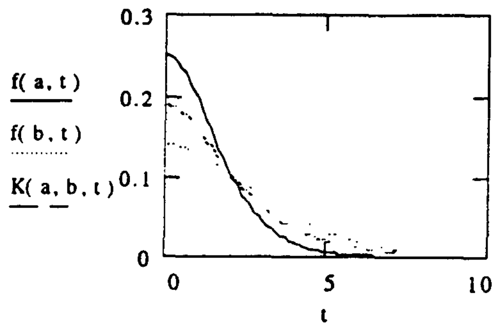
$a := 1$ $b := .55$ Select the scale factors.

$$C(y) := \int_{-20}^{20} K(a,b,y-x)\cdot f(a,x)\,dx$$

Define the convolution integral.

Check the functions:

$$t := 0, 0.1 .. \frac{4}{b}$$



Check normalization:

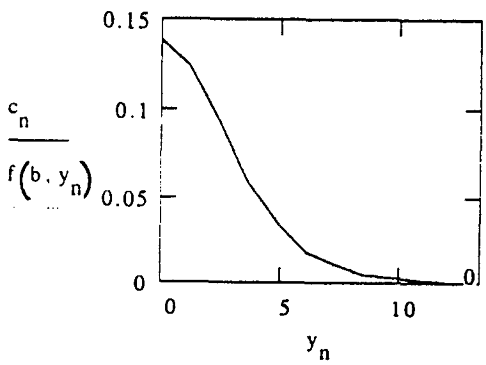
$$\int_{-20}^{20} f(a,x)\,dx = 1$$

$$\int_{-20}^{20} f(b,x)\,dx = 1$$

$$\int_{-20}^{20} K(a,b,x)\,dx = 1$$

Do the computation for a set of N points.

$N := 12$ $n := 0 .. N-1$ $y_n := \frac{4}{b}\cdot\frac{2\cdot n}{N}$ $c_n := C(y_n)$



The result should be identical to the scaled function.

Rewrite functions to parameterize by standard deviation. Check relation among widths.

Logistic PDF :

$$s := \sqrt{\int_{-20}^{20} x^2 \cdot f(1, x)\, dx} \qquad s = 1.814$$

$$g(\sigma, x) := f\left(\frac{s}{\sigma}, x\right)$$

$$L(\sigma, \tau, x) := K\left(\frac{s}{\sigma}, \frac{s}{\tau}, x\right)$$

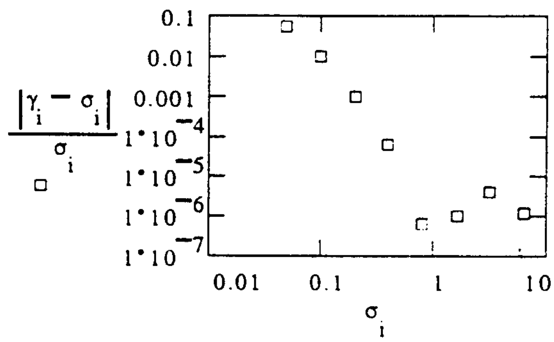
$$i := 0 .. 7$$

$$\sigma_i := 0.05 \cdot 2^i$$

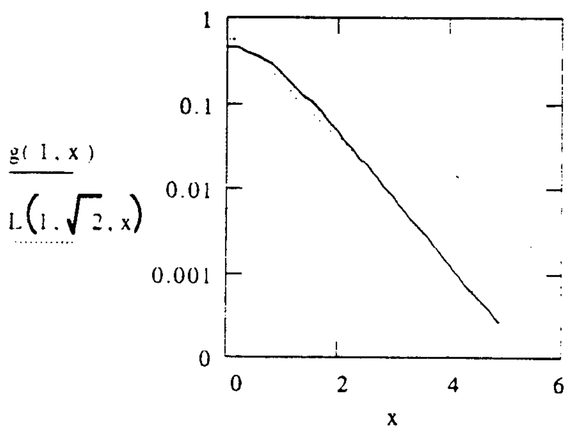
$$\tau_i := \sqrt{1 + (\sigma_i)^2}$$

$$\gamma_i := \sqrt{\int_{-10 \cdot \tau_i}^{10 \cdot \tau_i} x^2 \cdot L(1, \tau_i, x)\, dx}$$

To good accuracy, the widths add in quadrature. Works best for = input widths.



$$x := 0, 0.1 .. 4.9$$



These two functions have the same width.
The kernel has a sharper peak and a longer tail than the logistic pdf.

Look at the variation in shape of the kernel function as the parameters are varied subject to the constraiunt of constant width...

$$KW(\sigma, b, x) := K\left(\sqrt{\frac{b^2 \cdot s^2}{s^2 - b^2 \cdot \sigma^2}}, b, x\right)$$

$\sigma := 1$

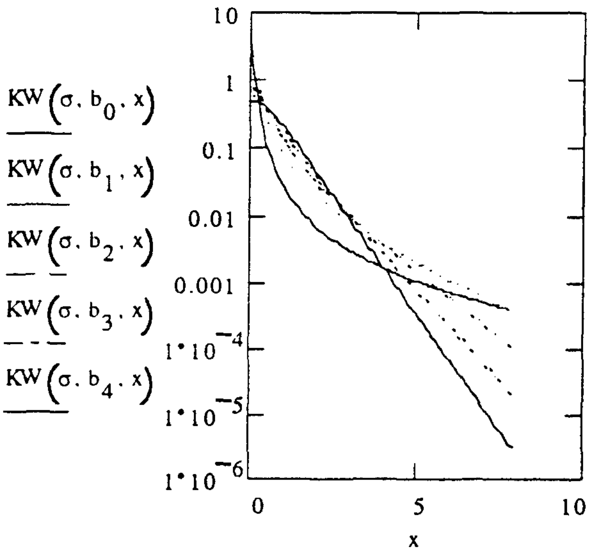
$N := 5$

$m := 0 .. N - 1$

$b_m := \frac{m + 0.5}{N} \cdot \frac{s}{\sigma}$

Argument of K must be real, so there is an upper limit on b.

$x := 0, 0.1 .. 7.9$



file cdf.mcd Jim Silk
Closed form of the cumulative distribution function for the kernel that connects scaled logistic functions.
The kernel K is defined in the file conv.mcd.
Our CDF is defined as follows:

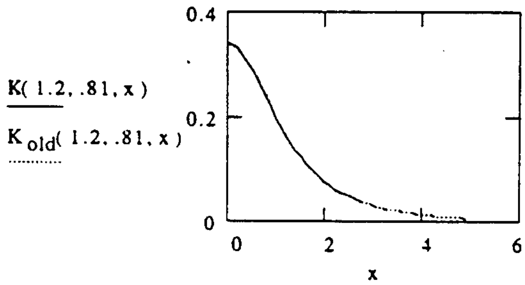
$$\mathrm{CDF}(a,b,x) := \int_{-\infty}^{x} K(a,b,x')\,dx'$$

Equivalent definitions of the kernel function: Result is from Bateman's Tables of Integral Transforms, A. Erdelyi, Volume 2 page 30 equations (2) and (6).

$$h(r,z) := \frac{1}{2\cdot\pi\cdot r}\cdot\frac{\sin(\pi\cdot r)}{\cosh(z)+\cos(\pi\cdot r)}$$

$$K(a,b,x) := b\cdot h\left(\frac{b}{a}, b\cdot x\right) \qquad K_{old}(a,b,x) := \frac{a}{2\cdot\pi}\cdot\frac{\sin\left(\pi\cdot\frac{b}{a}\right)}{\cosh(b\cdot x)+\cos\left(\pi\cdot\frac{b}{a}\right)}$$

$x := 0, 0.1 .. 4.9$



Check equivalence.

$$H_I(r,x) := \int_{-20}^{x} h(r,x')\,dx'$$

The integral form (definition) of the CDF.

$$H(r,x) := \frac{1}{\pi\cdot r}\cdot\mathrm{atan}\left(\frac{\sin(\pi\cdot r)}{\cos(\pi\cdot r)+\exp(-x)}\right)$$

This form has branch cut problems when the sine is negative, so use angle function.

$$H(r,x) := \frac{1}{\pi\cdot r}\cdot\mathrm{angle}(\cos(\pi\cdot r)+\exp(-x), \sin(\pi\cdot r))$$

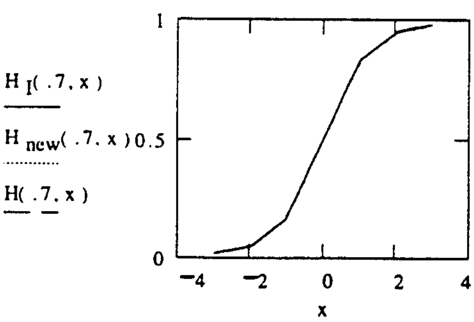
Easy to check that: $h(r,x) := \frac{d}{dx}H(r,x)$

$$H_{new}(r,x) := \frac{1}{2}+\frac{1}{\pi\cdot r}\cdot\mathrm{atan}\left(\tanh\left(\frac{x}{2}\right)\cdot\tan\left(\frac{\pi\cdot r}{2}\right)\right)$$

This form is just as good; see Gradshteyn 2.444.1.

$x := -3 .. 3$

$$\mathrm{CDF}(a,b,x) := H\left(\frac{b}{a}, b\cdot x\right)$$



Closed forms both agree with numerical integration.

APPENDIX B

Matlab™ CODE USED TO PERFORM UNCONSTRAINED FITS

APPENDIX B
Matlab™ CODE USED TO PERFORM UNCONSTRAINED FITS

```
load b1mat
p1=b1mat(1:69,:);
p2=b1mat(69+1:2*69,:);
p3=b1mat(2*69+1:3*69,:);
p4=b1mat(3*69+1:275,:);
p=p1;                                   %Select clutter
po=mean(p)';
pt=mean(p')';
[nt,no]=size(p);
lo=log((0.001*ones(size(po))+po)./(1.002*ones(size(po))-po));
lt=log((0.001*ones(size(pt))+pt)./(1.002*ones(size(pt))-pt));
l=[lo;lt];
l=[1.925*lo-2.613*(ones(size(lo)));1.366*lt]; %From 2-par fit
n_rep=199;
clear tmp;
clear tmp0;
tmp=l;
tmp0=tmp;
llrt=llr(l,p);                          %See function def. below.
for i=1:n_rep;
ld=demean(tmp(:,i:i),no,nt);            %See function def. below.
l=newton(ld,po,pt);                     %See function def. below.
tmp(:,i+1:i+1)=l;
tmp0(:,i+1:i+1)=tmp(:,i+1:i+1)-tmp(:,i:i);
llrt(i+1)=llr(l,p);
end;
s1=tmp(1:no,:)';
t1=tmp(no+1:no+nt,:)';
llr1=llrt';
save s1.dat s1 -ascii -tabs
save t1.dat t1 -ascii -tabs
save llr1.dat llr1 -ascii -tabs
```

```
function lnew=newton(l,po,pt);
%
no=size(po,1);
nt=size(pt,1);
lo=l(1:no);
lt=l(no+1:no+nt);
tmp=exp(-lt)*exp(-lo');
q=1./(ones(size(tmp))+tmp);
qo=mean(q)';
qt=mean(q')';
fd1=[po-qo;pt-qt];
sd0=q.*(ones(size(q))-q);
sqo=sum(sd0)';
sqt=sum(sd0')';
sd1=-[diag(sqo) sd0';sd0 diag(sqt)];
sd=sd1(2:nt+no,2:nt+no);
fd=fd1(2:nt+no);
l1=l(2:nt+no);
lnew1=l1-inv(sd)*fd;
lnew=[l(1:1);lnew1];


function ld=demean(l,no,nt);
%
lo=l(1:no);
lt=l(no+1:no+nt);
m=mean(lo);
ld=[lo-m*ones(size(lo));lt+m*ones(size(lt))];
```

APPENDIX C

Mathcad™ SCRIPT USED TO COMPUTE TWO PARAMETER FITS

New Obs • Tgt file.
Calculates two predictors from observer and target probability estimates.
Then do logistic fit to optimize the combination.

$PP := READPRN(matrix)$ $\qquad rows(PP) = 275$

$N_{rows} := rows(PP) \quad N_{cols} := cols(PP) \qquad i := 0..69-1 \qquad j := 0..N_{cols}-1$

Do Clutter Class #1: $P_{i,j} := PP_{i+0\cdot 69,\,j}$

$N_{rows} := rows(P) \quad N_{cols} := cols(P) \qquad N_{rows} = 69 \qquad N_{cols} = 22$

$i := 0..N_{rows}-1 \quad j := 0..N_{cols}-1$

$N_{obs_i} := \sum_j if(P_{i,j}<0,0,1) \qquad N_{obs_0} = 22$

$N_{tgt_j} := \sum_i if(P_{i,j}<0,0,1) \qquad N_{tgt_0} = 69$

$N_{net} := \sum_i \left(\sum_j if(P_{i,j}<0,0,1)\right) \qquad N_{net} = 1.518\cdot 10^3$

$P_{tgt_i} := \frac{1}{N_{obs_i}}\cdot\sum_j if(P_{i,j}<0,0,P_{i,j}) \qquad P_{obs_j} := \frac{1}{N_{tgt_j}}\cdot\sum_i if(P_{i,j}<0,0,P_{i,j})$

$P_{net} := \frac{1}{N_{net}}\cdot\left[\sum_i\left(\sum_j if(P_{i,j}<0,0,P_{i,j})\right)\right] \qquad P_{net} = 0.769$

$limit(x) := if\left(|x| < .99, x, .99\cdot\frac{x}{|x|}\right) \qquad logodds(x) := \ln\left(\frac{x}{1-x}\right)$

$lo(x) := 2\cdot atanh(limit(2\cdot x - 1)) \qquad logit(x) := 0.5\cdot(tanh(0.5\cdot x)+1)$

$DATA_{i+j\cdot N_{rows},\,0} := P_{i,j} \qquad k := 0..N_{rows}\cdot N_{cols}-1$

$DATA_{i+j\cdot N_{rows},\,1} := lo\left(P_{obs_j}\right) \qquad k_{max} := \sum_k if(DATA_{k,0}<0,0,1)$

$DATA_{i+j\cdot N_{rows},\,2} := lo\left(P_{tgt_i}\right)$

$DDATA := csort(DATA, 0) \qquad sbar := mean(DDATA^{<1>}) \qquad mean\left(\overrightarrow{lo(P_{tgt})}\right) = 1.818$

$k := 0..k_{max}-1 \qquad sbar = 1.347 \qquad stdev\left(\overrightarrow{lo(P_{tgt})}\right) = 2.239$

$idat_{k,1} := DDATA_{k,1} - sbar \qquad mean\left(\overrightarrow{lo(P_{obs})}\right) = 1.347$

$idat_{k,0} := DDATA_{k,2} \qquad stdev\left(\overrightarrow{lo(P_{obs})}\right) = 0.772$

$ddat_k := DDATA_{k,0}$

$ndat := rows(idat)$ $\quad ndat = 1.518 \cdot 10^3$ $\quad N_{obs} := 1$

$npar := cols(idat) - 1$ $\quad npar = 1$ $\quad P := ddat$

$i := 0 .. ndat - 1$ $\quad idat_{i,0} := 1$ $\quad P_{bar} := \frac{\sum P}{ndat}$ $\quad P_{bar} = 0.769$

$k := 0 .. npar$ $\quad \beta_k := 0$ $\quad unit_k := 1$

$logits(\beta) := idat \cdot \beta$

$pred(lgt) := \overrightarrow{\left[\frac{1}{2} \cdot \left(1 + \tanh\left(\frac{lgt}{2}\right)\right)\right]}$

$Q(\beta) := pred(logits(\beta))$

$Q2(q) := \overrightarrow{(q \cdot (1 - q))}$

$xtnd(q) := q \cdot unit^T$

$Idt(QQ) := \overrightarrow{(idat \cdot QQ)}$

$Idat(\beta) := Idt(xtnd(Q2(Q(\beta))))$

$LLR0(q) := N_{obs} \cdot \sum \overrightarrow{(P \cdot \ln(q) + (1 - P) \cdot \ln(1 - q))}$

$LLR(\beta) := LLR0(Q(\beta))$

$FD(\beta) := N_{obs} \cdot \left[idat^T \cdot (P - Q(\beta))\right]$

$SD(\beta) := -N_{obs} \cdot \left(idat^T \cdot Idat(\beta)\right)$

$SAT := LLR0(0.0001 + .9998 \cdot P)$ $\quad SAT = -0.152$

$$\beta := \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$B^{<0>} := \beta \quad N := 4 \quad n := 0..N$

$B^{<n+1>} := B^{<n>} - SD(B^{<n>})^{-1} \cdot FD(B^{<n>})$

$$B^T = \begin{bmatrix} 1 & 2 \\ 1.239 & 1.83 \\ 1.345 & 1.912 \\ 1.357 & 1.924 \\ 1.357 & 1.924 \\ 1.357 & 1.924 \end{bmatrix}$$

$\beta := B^{<N+1>} \quad LLR(\beta) = -403.4$

$D := 2 \cdot (SAT - LLR(\beta)) \quad D = 806.496$

$y_2(x, n_D) := \sqrt{2 \cdot x} - \sqrt{2 \cdot n_D - 1}$

$cnorm(-y_2(D, ndat - npar - 1)) = 1$

$y_2(D, ndat - npar - 1) = -14.892$

$$FD(\beta) = \begin{pmatrix} 9.886 \cdot 10^{-13} \\ -1.036 \cdot 10^{-13} \end{pmatrix} \qquad SD(\beta) = \begin{pmatrix} -287.734 & 89.387 \\ 89.387 & -74.993 \end{pmatrix}$$

$$\beta = \begin{pmatrix} 1.357 \\ 1.924 \end{pmatrix}$$

$I := -SD(\beta) \quad \Sigma := I^{-1}$

$$\Sigma = \begin{pmatrix} 0.006 & 0.007 \\ 0.007 & 0.021 \end{pmatrix}$$

$\lambda := eigenvals(\Sigma) \quad s^{<k>} := eigenvec(\Sigma, \lambda_k) \quad SE_k := \sqrt{\Sigma_{k,k}} \quad Wald := \overrightarrow{\frac{\beta}{SE}}$

β_k	SE_k	$Wald_k$
1.357	0.074	18.271
1.924	0.146	13.225

$$\lambda = \begin{pmatrix} 0.003 \\ 0.024 \end{pmatrix} \qquad s = \begin{pmatrix} -0.94 & 0.342 \\ 0.342 & 0.94 \end{pmatrix}$$

$$s^T \cdot s = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad s^T \cdot \Sigma \cdot s = \begin{pmatrix} 0.003 & 0 \\ 0 & 0.024 \end{pmatrix}$$

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
April 1997

3. REPORT TYPE AND DATES COVERED
Final—October 1994–June 1995

4. TITLE AND SUBTITLE
Modeling the Observer in Target Acquisition

5. FUNDING NUMBERS
DASW01 94 C 0054
DARPA Assignment A-181

6. AUTHOR(S)
James D. Silk

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Institute for Defense Analyses
1801 N. Beauregard St.
Alexandria, VA 22311-1772

8. PERFORMING ORGANIZATION REPORT NUMBER
IDA Paper P-3102

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Research Projects Agency
3701 N. Fairfax Drive
Arlington, Virginia 22203-1714

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 180 words)*

Models of target acquisition by human observers are probabilistic. Two reasons for the indeterminate nature of the models are (1) the observer ensemble includes individuals with varying levels of competence in target acquisition, and (2) the target acquisition process is itself (at least in part) a random process. Each of the two commonly used combat simulations, CASTFOREM and JANUS, assumes that (a different) one of these reasons completely accounts for the indeterminacy, and ignores the other. Although the distinction is irrelevant for the simple one-on-one case, it has recently been shown that the choice can profoundly affect the outcomes of more realistic many-on-many engagements. We propose a model in which target detection probability is a function of both a target statistic *and an observer statistic.* Our analysis of recently compiled observer test data validates the model and provides it with the correct quantitative balance between variations among observers and the inherently stochastic component of target acquisition. We provide a working simulation that precisely represents our model. The algorithm is suitable for insertion into the wargames; it builds on the approaches of its predecessors but requires only a small increase in complexity.

14. SUBJECT TERMS
target acquisition model, force models, wargames, human factors, observer population statistics

15. NUMBER OF PAGES
50

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR