VALIDATION OF A TARGET RANGE DECISION AID

(THIS PAPER IS UNCLASSIFIED)

Marvin S. Cohen
Decision Science Consortium, Inc.

ABSTRACT

In passive target ranging aboard a nuclear attack submarine several analytical procedures exist for estimating a single variable, and the estimates which are obtained at a given time frequently do not agree. Typically, the Commanding Officer focuses on one or two of these estimates, thus failing to utilize additional information that is, in principle, available. To address this problem, an interactive Bayesian decision aid computes a single pooled estimate of target range, together with an assessment of its probable accuracy. The aid can operate either automatically or by incorporating subjective judgments concerning the validity of the estimates being pooled.

Quantitative tests with pre-recorded at-sea data (Rangex 1-78 and 1-79) evaluated the aid with objective inputs, subjective inputs, and a mix of the two. It was found that estimates computed by the pooling aid in all three of these modes were consistently more accurate than EAST, MATE, and Ekelund ranging procedures and the Command staff's unaided best guess as to target range. Maximum accuracy was obtained when Command staff subjective judgments were combined with objectively assessed parameters.

Formal analysis of the sensitivity of pooling accuracy to errors in inputs suggests that the advantage of pooling over reliance on a single estimate will be extremely robust over a wide range of conditions.

INTRODUCTION

The effectiveness of command decision aids may be enhanced by a growing sense of their limitations. Since "objective" methods can seldom address more than part of a complex problem, prescriptive aids are appropriately regarded not as "know-it-alls" but as fallible "advisors" (Patterson el al., 1981). The decision maker's own experience may be the best (or only) source of relevant information in some matters (e.g., uncertainty about the intentions of a military foe, or the competence of an operator supplying range estimates), while an exclusively "factual" approach could be fatally incomplete.

To earn the Commander's trust, a decision aid must display evidence as well as conclusions, accommodating his requests for information at any level of detail (e.g., from raw bearing data to target range estimates to "probability of kill"). Such aids will allow the Commander to interpose his own assessments in addition to or in place of default values at any level. But they will rapidly and systematically integrate subjective inputs with the objective data that are retained.

In this report we test the feasibility of a probabilistic aid which helps the Commander of a nuclear attack submarine evaluate noisy and sometimes inconsistent information about target range. Within this special context, we test the hypothesis that (quite aside from any impact on user acceptance) incorporation of Command staff judgments in addition to objective data can enhance the quantitative accuracy of decision aid output.[1]

Rationale

The requirement of covertness in submarine-based antisubmarine warfare (ASW) imposes constraints on the quality of information available to the attacking platform. For example, data provided by passive sensors (which do not alert the enemy) is usually less good than data that could be obtained by actively pinging. Numerous techniques are available for extracting information regarding the location and motion of a target - based on different aspects of the data (e.g., bearing, intensity, angle between direct and reflected sound paths) and using different analytical tools and assumptions. Typically (since their sources of error are both pronounced and different) they produce quite diverse estimates of target range. The Commander, nonetheless, must use these estimates to make critical decisions regarding approach maneuvers and time of fire - decisions which are heavily dependent on his assessment of range-related risks (e.g., counterdetection) and opportunities (e.g., probability of a hit).

A current tool for informally integrating various range solutions is the Time/Range (T/R) Plot, on which range estimates labelled by source are plotted (usually by hand) against time. The T/R Plot falls short of what could be achieved by more systematically integrating data about target range:

- It provides no assurance that all the information in the various ranging techniques is utilized. Confronted with a divergent set of estimates, the CO is likely either to suspend judgment about range or to focus on only one or two of the available estimates.

VALIDATION OF A TARGET RANGE DECISION AID

Marvin S. Cohen

- There is no explicit measure of overall range solution quality. As a result, the CO may postpone attack in order to refine localization more than it needs to be refined.

- Range information is not integrated with information about threat and own ship capabilities, to support critical command decisions. Inappropriate use of range information could lead to a premature attack (which both alerts the enemy and misses) or to unnecessary delay (increasing the risk of counterdetection or loss of opportunity).

Ideally, of course, it would be desirable to have a single accurate procedure for estimating target range: an exact formula applied to exactly measured inputs. Unfortunately, despite recent and expected improvements in sensor systems and ranging algorithms, the ideal is not realistically attainable. Any passive localization technique, no matter how good, will almost certainly be subject both to significant error in the measurement of inputs (e.g., bearings, thermal layer depth) and to occasional failure of assumptions (e.g., about target maneuvers) upon which its validity depends.

Because no individual technique is free of uncertainty, a complementary strategy suggests itself: the design of a framework to exploit all the information (however uncertain) in a plurality of ranging techniques, to assess the amount of uncertainty that remains, and to combine that information probabilistically with relevant prior information about weapon and sensor capabilities. Such a procedure is largely independent of the details of the ranging techniques being pooled, and is by no means inconsistent with efforts to improve one or another of the individual techniques. Yet as long as no one technique exhausts the relevant evidence, dramatic improvements in target localization accuracy (and sharper inferences about combat-critical events like counterdetection and target kill) may be achieved, in principle, within such a framework. The pooled estimate is based on a larger fund of data than any particular solution, even the one that is, on the average, best.

A Range Pooling Technique

The basis for the approach adopted here is an application of Bayesian probability theory, by means of which the evidential impact of each ranging technique is quanitifed in the context of other techniques and used to adjust a prior range estimate. The outcome of a series of such adjustments is the overall pooled solution.

If each range estimate takes the form of a probability distribution f_i over target range R, a single pooled distribution can be computed which reflects their total evidential impact:

$$F(R|\underline{f}) \propto F(\underline{f}|R)F(R) \tag{1}$$

$$= F(f_n|f_1,f_2,\ldots f_{n-1},R)\ldots F(f_2|f_1,R)F(f_1|R)F(R),$$

where $F(\cdot)$ denotes probabilities assessed by the pooling aid or its user.

It might be possible to pool range solutions directly using the above formula. However, by making a few additional assumptions, computations will be speeded and simplified, and inputs will be transformed into a more readily understood form, so that users can evaluate such inputs and (if they choose) subjectively revise them.

Suppose that information about each f_i consists of a mean E_i and variance V_i. Then in place of equation (1), we have:

$$\begin{aligned} F(R|\underline{E},\underline{V}) &\propto F(\underline{E},\underline{V}|R)F(R) \\ &= F(\underline{E}|\underline{V},R)F(\underline{V}|R)F(R). \end{aligned} \tag{2}$$

We now assume:

(a) $F(\underline{V}|R)$ is independent of R.

(b) $F(\underline{E}|\underline{V},R)$ is multivariate normal, with means $R+B_i$, variances V_i, and covariances Cov_{ij}.

(c) $F(R)$ is relatively invariant with R in the region of interest.

For additional discussion of some of these issues, see Lindley, Tversky, and Brown, 1979; Morris, 1977; Winkler, 1981; Cohen and Brown, 1980. Assumptions (a) and (b) are almost certainly false; we will see, however, that the impact of these approximations on pooling accuracy is likely to be quite small.

As a result of these assumptions, the pooled range estimate can be computed as a weighted average of the original estimates, after correcting for bias:

$$E_{1,2} = W(E_1-B_1) + (1-W)(E_2-B_2) \tag{3}$$

where

$$W = \frac{p-\rho\sqrt{p}}{1+p-2\rho\sqrt{p}} \tag{4}$$

and p is relative precision (V_2/V_1). The expected variance of the pooled solution is a function of the variances and covariances of the estimates being pooled:

$$V_{1,2} = \frac{V_1V_2-Cov_{1,2}^{\;2}}{V_1+V_2-2Cov_{1,2}} \tag{5}$$

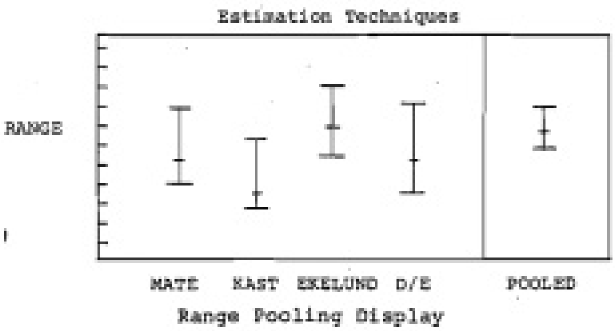
Default values for pooling aid parameters (biases, variances, and covariances) will be estimated beforehand (e.g., from exercise data, as described in the Method Section) and stored for automatic use, subject to the Commander's adjustment. These values will be conditional on possible scenarios - e.g., thermal conditions, number of data points, maneuver geometry, or estimated range. The range estimates themselves will either be supplied automatically (through the Fire Control System) or be manually input (e.g., estimates based on manual plots). The output of the algorithm is a single pooled range estimate and an assessment of its error variance, expressible as an interval within which the true range is expected to fall with a specified probability (e.g., 95%).

VALIDATION OF A TARGET RANGE DECISION AID
Marvin S. Cohen

Displays

Figure 1 is a display of the pooled solution together with inputs upon which it is based. Expected values and 95% intervals of uncertainty are displayed for each available ranging technique, as well as for the pooled estimate. The user can adjust default values for inputs and observe the implications of his revision for the pooled estimate. (The default values, however, continue to be stored and available for display.)

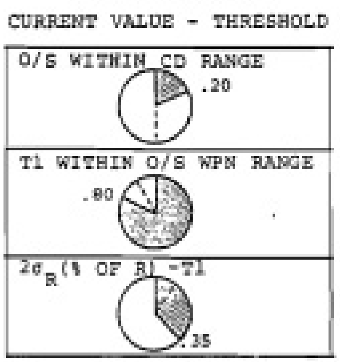
UNCLASSIFIED



Range Pooling Display
Figure 1.

In Figure 2 the Commander has requested current status displays for three range-related variables: the probability of being within the target's counter-detection range (.20), the probability of having the target within own ship weapon range (.80), and the 95% interval of uncertainty as a percentage of the pooled range estimate (±35%). The Commander can set a threshold on any of these variables, as indicated by the dotted lines. When such a threshold is crossed, the Commander will be alerted.

UNCLASSIFIED



Alerting Display
Figure 2.

Other displays involving target range will be required for other purposes: e.g., a Time/Range plot of pooled range estimates (with uncertainty intervals), and a geographic display utilizing ellipses to represent target location uncertainty. Fuller discussion of such display possibilities within the context of a general attack planning decision aid is contained in Cohen et al. (1982a, 1982b) and Cohen (1982).

METHOD

Whether pooling improves range accuracy (and whether it assesses accuracy correctly) will be decided, ultimately, at sea. The present section reports the results of some preliminary quantitative tests - simulating the operation of the pooling aid and its use by the Commander with pre-recorded data from at-sea exercises.

Three different modes of operation of the pooling aid were put to test:

- Automatic mode: Pooling takes place without user interaction, employing default parameters derived from exercises.
- Previously adjusted weights: Prior to an engagement, the user has inserted his own estimates of the relative validity of different ranging techniques in place of default values.
- On-the-spot adjustments: During an engagement, the user arrives at his own range estimates (or assessments of relative validity) and adjusts the automatically provided values accordingly.

Each of these modes was compared to three specific ranging techniques (MATE, KAST, and Ekelund) and to the unaided best guess of the Command staff (recorded onboard as the "system solution").

Data

Data utilized in this study came from on-range baseline runs by two U.S. nuclear attack submarines (to be designated SSN X and SSN Y) in Rangex 1-78 and Rangex 1-79, respectively. Runs were conducted on the weapons range of the Atlantic Undersea Test and Evaluation Center (AUTEC). All runs contained one target maneuver and a mixture of opening and closing geometries, with ranges typical of MK 48 torpedo deployment. Both exercises included variations in leg length (long and short) and target signal-to-noise ratio (SNR) (high and low).

A total of 200 data points were extracted, one at the end of each own ship leg, up to the target maneuver. The following information relating to target range was extracted, when available, for each data point: MATE solution, KAST solution, Ekelund solution, system solution, and reconstructed range. All but the last of these were extracted from on-board

VALIDATION OF A TARGET RANGE DECISION AID

Marvin S. Cohen

records - either manual logs or the automatic Data Gathering System (DGS) associated with the MK 117 or MK 113 Mod 10 Fire Control System.

MATE, KAST, and Ekelund are all "bearings only" techniques whose quality depends directly on bearing measurement error, the number of data points, the geometry of own ship maneuvers, and the assumption that the target has not maneuvered. MATE (Manual Adaptive Target-Motion-Analysis Evaluation) is an interactive method in which an operator inputs hypotheses about target range, course, and speed and observes a computer display of the differences between observed and predicted bearings. KAST (Kalman Automatic Sequential Target-Motion-Analysis) produces automatic solutions based on assumptions about the normality of errors and their relation to SNR. Ekelund is more approximate method for obtaining a quick range estimate based on changes in the rate of bearing change as a result of an own ship maneuver. The "system solution" is continuously updated in the Fire Control System by the Command staff and serves as the source of weapon settings in an attack. It represents, in principle, the Command staff's current best guess as to target range. (No system solution was identified for Rangex 1-79, although the MATE solution may in fact have been utilized in this way.) Reconstructed ranges are based on data from ocean-bottom sensors and approximate reasonably closely the true values.

<u>Test Conditions</u>

We have utilized this data in several quite different ways:

(1) Default weights. First, exercise data was used to estimate default parameters for the pooling aid by means of standard formulae. The error for each solution at each data point is the difference between the solution estimate and the reconstructed range. Default parameter values for each solution technique are statistical functions of these errors: average error (bias), variance of errors, and pairwise covariances of errors. Weights based on these parameters were used to pool pairs of estimates: MATE/KAST, MATE/Ekelund, and Ekelund/KAST. To pool all three solutions, an iteration was required: Error statistics for the pooled solution involving MATE and KAST were computed, and then this solution (MATE + KAST) was pooled with Ekelund. (Henceforth, the "default pooled solution" for any data point will refer to the solution which includes the maximum possible number of pooling iterations.)

In the absence of a "live" Command staff, we have employed exercise data (in a more artificial way) to simulate two sorts of user interaction with the pooling aid:

(2) Subjective weights. The Command staff relies, to varying degrees, on specific ranging techniques in arriving at its own best guess as to target range (the system solution). To simulate subjective judgments of estimate validity prior to an actual engagement, this reliance was modeled. Weights which best predicted the system solution were fit to each estimate by least squares. The effect of substituting these weights into the pooling algorithm, in place of the default values, is presumed to reflect at least roughly the outcome of pooling with weights subjectively assessed by the Command staff.

(3) Subjective adjustments of default estimates. During an engagement, adjustments by the Command staff of default pooled range solutions would presumably be in the direction of its own best guess as to target range. To simulate an adjustment of this sort, the system solution was pooled with the default pooled solution. Weights used for this higher-order pooling determine the magnitude of the adjustment, and reflect the relative validity and correlation of the system solution and default pooled solution.

(4) Cross-validation. To achieve a limited degree of cross-validation, the total sample was randomly divided into two parts - a "first sample" for the purpose of estimating parameters (i.e., default weights, subjective weights, and higher-order pooling weights), and a "second sample" for testing the performance of the aid. The gain in accuracy due to pooling with default inputs was also examined within divisions of the data according to platform (SSN X/SSN Y), SNR (high/low), and leg length (short/long).

The primary measure of accuracy utilized is mean absolute error (MAE): the average difference between a range estimate and the true range, without regard to the direction of the difference. (However, results in terms of the standard deviation of errors are reported in Cohen (1982) and are quite comparable.) Only relative measures of solution accuracy will be given: The error measure for a particular ranging technique will be reported as a proportion of the same error measure for the default pooled solution in the same data sample, values greater than one thus represent an advantage for pooling with default weights in that data. (For example, using hypothetical numbers, suppose the MAE for MATE were 300 yards, and the MAE for the default pooled solution were 250 yards. Then the relative MAE for MATE would be reported as 300/250 = 1.20. This represents a 20% increase in error due to using MATE as opposed to the pooling aid.)

RESULTS

<u>Accuracy of Default Pooled Solution</u>

Pooling parameters estimated from the first data sample are given in Table 1. In this data MATE was more precise than both KAST and Ekelund; KAST was more precise than Ekelund, as was the pooled solution involving MATE and KAST. The only correlation of moderate size is between errors in KAST and MATE.

UNCLASSIFIED

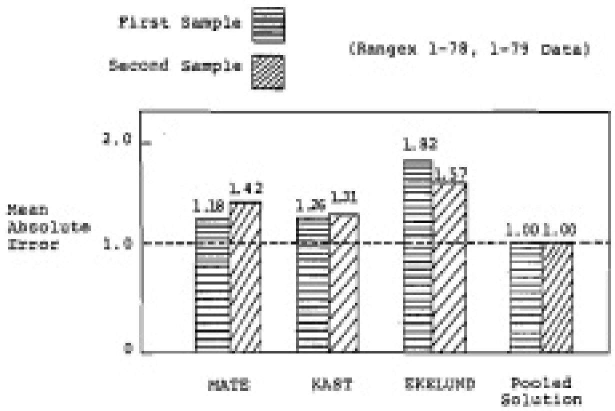
		Weights Based on:	
Solutions Pooled	Default Weights	Relative Precision	Correlation
MATE KAST	.60 .40	1.3	ρ = .34
MATE EKELUND	.72 .28	2.5	ρ = -.02
KAST EKELUND	.66 .34	1.9	ρ = .02
MATE-KAST EKELUND	.77 .23	3.2	ρ = -.02

Default Parameters Used in Pooling
Table 1.

VALIDATION OF A TARGET RANGE DECISION AID
Marvin S. Cohen

Figure 3 compares the pooled solution to the specific techniques being pooled in terms of mean absolute error. Pooled solution results in the two samples have each been normalized to 1.00 (with no implication that MAE for pooling was actually the same in the two samples). In the First Sample, the ratio of the MAE for KAST to the MAE for the pooling aid was 1.26 - for an increase of 26% in absolute error due to using KAST. In the Second Sample, the increase was quite comparable, at 31%. Results for MATE and Ekelund, for both first and second samples, also show larger mean absolute errors than with pooling.

UNCLASSIFIED



	MATE	KAST	EKELUND	Pooled Solution
Sample Size	82, 97	88, 99	82, 95	90, 100
Statistical Significance of Comparison With Pooled Solution	<.001	<.001	<.001	

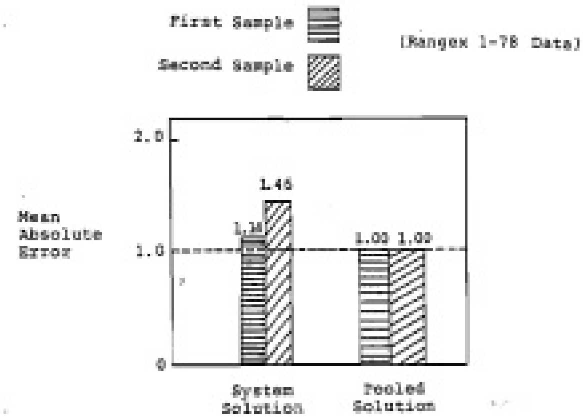
Figure 3. Ratio of Mean Absolute Error (MAE) for Current Ranging Techniques to MAE for Pooling Aid. Significance level is based on t-test, utilizing second sample data only.

Figure 4 compares pooling with the system solution. In the first and second samples, use of the system solution would have increased MAE by 14% and 46%, respectively, over use of the pooled solution.

Improved accuracy generally counts the most when quite large range errors are avoided - since small errors could be overcome by the weapon's search capability. Thus, Table 2 focuses on the contribution of pooling to the reduction of <u>large</u> range errors - i.e., errors in either direction which are greater than 20% of the true range. It compares MATE, KAST, Ekelund, and the system solution to the pooling technique in terms of the proportion of the time an error is large. The proportion of large errors is 55% and 17% larger for the system solution than for the pooled solution, in samples one and two, respectively. The figures are 18% and 72% for MATE; and so on.

Does the effectiveness of the range pooling aid depend on which boat utilizes it? Figure 5 breaks down the results for mean absolute error from both samples according to whether SSN X or SSN Y is involved. The same default parameters were used for both platforms. The improvement due to the pooling is quite comparable in both cases. Similarly, the contribution of pooling to accuracy was unaffected by breakdown of the data according to leg length and signal-to-noise ratio (Cohen, 1982).

UNCLASSIFIED



	System Solution	Pooled Solution
Sample Size	50, 57	50, 55
Statistical Significance of Comparison (t-test)	.085	

Figure 4. Ratio of Mean Absolute Error (MAE) for System Solution to MAE for Pooling Aid

UNCLASSIFIED

		First Sample	Second Sample
Ranging Techniques	MATE	1.18	1.72
	KAST	1.14	1.33
	EKELUND	1.95	1.72
System Solution		1.55	1.17
Pooling Aid		1.00	1.00

Table 2. Ratio of Proportion of Large Errors (>20% of true range) for Current Ranging Techniques to Large-Error Proportion for Range Pooling Aid

VALIDATION OF A TARGET RANGE DECISION AID

Marvin S. Cohen

UNCLASSIFIED

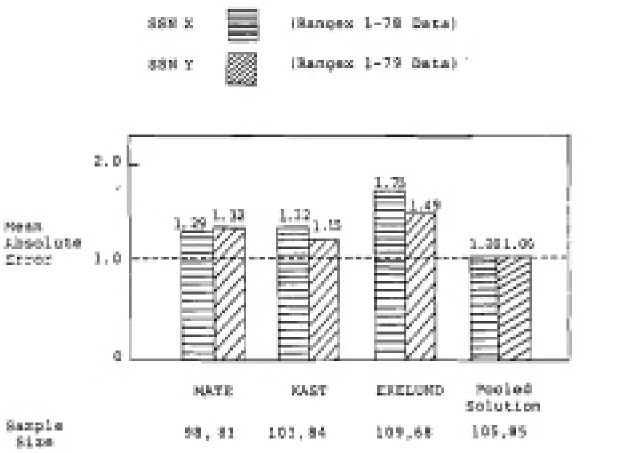


Figure 5. Ratio of Mean Absolute Error (MAE) for Current Ranging Techniques to MAE for Pooling Aid, Broken Down by Platform.

In sum, pooling with default parameters increased ranging accuracy in both the initial and the cross-validated data samples. Pooling was more accurate than any of the three individual ranging techniques (MATE, KAST, and Ekelund) and was also more accurate than the system solution, representing the Command staff's best guess as to target range. The superiority of the pooled solution occurred regardless of platform, SNR, or leg length.

Because of the relatively small number of data points, default pooling parameters were estimated from first sample data combining levels of leg length and SNR. We expect that pooling accuracy would have been even better had the weights been optimized within levels of such variables. But how much accuracy is lost by not doing so? Some perspective is provided by the sensitivity analysis to be reported in a later section. There is a rather large leeway for deviation of pooling weights from optimality, consistent with the superiority of pooling over reliance on any single estimate.

Accuracy of Pooling with Subjective Weights

We have evaluated the pooling aid in its "automatic" mode - utilizing default parameter values based on objective data. A critical feature of the aid, however, is its ability to incorporate subjective assessments by Command staff personnel whose knowledge of their situation may not be reflected in previous data.

Table 3 shows the least squares weights that describe the relation of the system solution to other range estimates. A higher weight for MATE than KAST means that MATE had a large influence on the system solution (i.e., Command staff range estimates) than KAST. There is, of course, no reason to suppose that these are the weights that would be supplied explicitly by the Command staff, or that would be derived from their direct judgments of validity and correlation. Nonetheless, it would be surprising if these weights did not reflect, at least in a gross way, Command staff judgments.

UNCLASSIFIED

Solutions Pooled	"Subjective" Weights Based on System Solution	Default Pooling Aid Weights
MATE	.78	.60
KAST	.22	.40
MATE-KAST	.84	.77
EKELUND	.16	.23

Table 3. Weights Derived By Modeling Command Staff

Comparison with the default pooling aid weights (from Table 1) shows no very large disparities. However, the Command staff weights tend to be more extreme than weights based on objective data. This may reflect less tendency to combine estimates and a greater tendency toward exclusive reliance on the more preferred solution.

What happens if range estimates are actually pooled using these implied subjective weights - rather than the weights assessed from objective data? Figure 6 provides the relevant comparisons: In both data samples, pooling with "subjective" weights was more accurate than the system solution itself. In the first sample, system solution MAE was 1.14 compared to 1.09 for subjective pooling; in the second sample, the difference is 1.46 to 1.18.

The advantage of pooling with subjective weights was not statistically significant. Then, too, it might seem odd that a model of the Command staff judgment process would be more accurate than the Command staff itself: presumably, the model is incomplete, since it omits other solutions of which the Command staff was aware. This result, however, corresponds to findings in a number of areas where simple models have outperformed the experts upon whom the models were based (e.g., Dawes, 1975, 1979). Such findings would be expected if the use of the model by the experts was disturbed by substantial random error - of which the model itself, of course, is free. Another possibility is that the experts are behaving quite differently from the model: e.g., rather than mentally combining estimates, the Command staff may be utilizing only a single technique on each occasion, perhaps shifting from one to another probabilistically. As we shall see, such a strategy would, under most circumstances, be expected to be less accurate than pooling. The validity of the weights used in pooling is less important, for accuracy, than the fact that some sort of pooling is taking place.

These findings have an interesting practical application. One way for the Command staff to improve ranging accuracy might be to assess pooling weights,

VALIDATION OF A TARGET RANGE DECISION AID
Marvin S. Cohen

UNCLASSIFIED

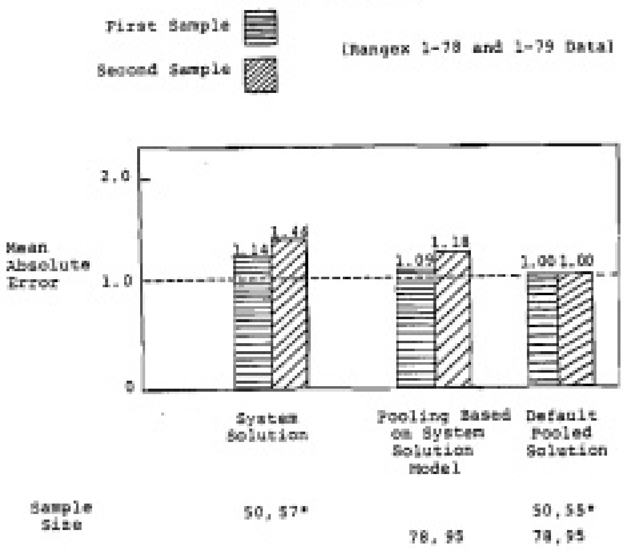


Figure 6. Ratio of Mean Absolute Error (MAE) for System Solution and for System Solution Model to MAE for Default Pooling Aid.

*=Rangex 1-78 data only

rather than range itself. (Alternatively, weights might be extracted, as they were here, from Command staff range estimates.) The <u>weights</u> would then be used by the pooling aid algorithm, to produce the "system solution."

Note, however, that this may not be optimal when suitable objective data are available. Figure 7 shows that pooling based on "subjective" weights was still not as accurate as pooling based on default values.

Accuracy of Subjective Adjustment of Default Estimate

The previous section has dealt with a special case of the range pooling aid: in which subjective estimates of weights entirely replace the default values obtained from objective data. More typically, one might imagine a compromise: adjustment of weights, or of range estimates, in the direction of - but not all the way to - subjectively preferred values.

Table 4 gives the parameters, estimated from the first data sample, which were utilized in the higher-order pooling of the system solution and the default pooled solution. The latter receives a larger weight than the system solution due to its greater accuracy. The correlation is negligible.

The result of pooling subjective and objective estimates is a solution which is (non-significantly) more accurate than either type of estimate alone. Figure 7 shows that the mean absolute error for the combined solution is 81% and 97% of the MAE for the default pooled solution, in the first and second samples, respectively.

Command staff judgment (the system solution) by itself is less accurate than the default pooled solution. Nonetheless, it may contain information not

UNCLASSIFIED

		Weights Based on:	
	Weights	Relative Precision	Correlation
System Solution	.35	.53	
			ρ = .04
Default Pooled Solution	.65		

Table 4. Parameters Used in Combining System Solution and Default Pooled Solution

UNCLASSIFIED

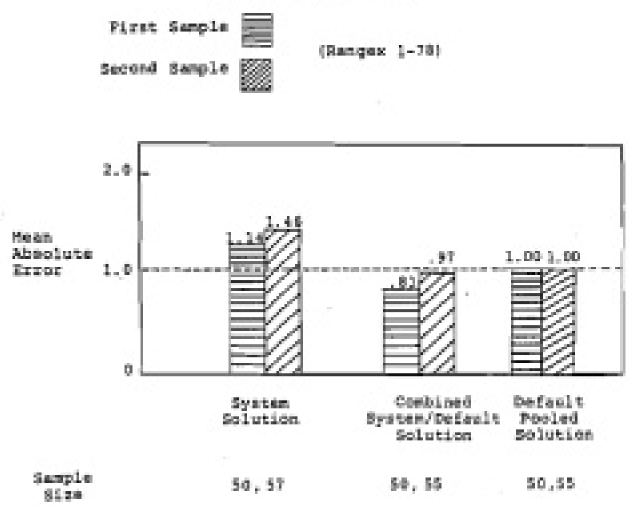


Figure 7. Ratio of Mean Absolute Error (MAE) for Combined Default/System Solution to MAE for Default Pooled Solution

captured by default parameters. Subjective adjustment of default solutions by Command personnel may be an easy and effective way to tap that information. (Alternatively, the Command staff and the pooling aid might work in parallel, arriving at separate estimates, and the results might then be formally pooled, as in the procedure tested here.)

To what extent does the advantage of combining objective and subjective information depend on the method of combination? In this test, we have loaded the dice in favor of such an advantage by formally pooling with weights based on first sample data. Informal subjective adjustments of default solutions would be unlikely to correspond to optimal results of higher-order pooling.

VALIDATION OF A TARGET RANGE DECISION AID
Marvin S. Cohen

Fortunately, the sensitivity analysis to be reported shortly suggests that the advantage of combining estimates is quite robust in the face of parameter-estimation errors. Even when the weights utilized in pooling deviate significantly from optimal weights, pooling is likely to outperform both of the estimates being pooled. Informal adjustments of the default solution would be expected to have a similar latitude. Note that if the alternative to informal adjustment were exclusive reliance on the less accurate estimate (in this case, the system solution), informal adjustment should produce improvements regardless of the degree of adjustment employed.

SENSITIVITY AND ROBUSTNESS

What happens if other than optimal parameters are used for pooling? As we have seen, the question is important for at least two reasons:

(1) In practice, a range pooling aid will apply default parameters in situations quite different from the ones in which they were estimated.

(2) Subjective adjustments of default parameters may be based on erroneous information.

How far off can parameters be - whether pooling operates automatically or interactively - without sacrificing the advantages of pooling?

The variance of a weighted average of two estimates, in relation to the variance of the more precise estimate (V_1), can be expressed in terms of p and ρ:

$$V_{1,2}(W)/V_1 = W^2 + p(1-W)^2 + 2\rho\sqrt{p}\,W(1-W). \tag{6}$$

Figure 8 plots this as a function of W for p=1.3 and ρ=.34. These are the actual values for pooling MATE and KAST obtained from the first sample (Table 1). Any choice of weights for pooling that are between 0 and 1 corresponds to a point on the solid curve; it predicts the variance of pooling with those weights under the specified real-world conditions, i.e., values of ρ and p.

The optimal (default) weight is W=.6, as given by formula (4), and corresponds to the low point of the function (i.e., formula (5)). The brackets in Figure 8 represent the maximum expected improvement in accuracy attainable by pooling, as compared with exclusive reliance on MATE, the better of the two estimates (E_1). Thus, pooled solution variance, with optimal weights, is 75% of the variance of E_1. Other data points shown correspond to exclusive use of MATE or KAST, pooling with equal weights for MATE and KAST, default weights with covariance set to zero, and Command staff subjective weights.

How far from optimal does W have to be for pooling to be no better than exclusive use of E_1? As Figure 8 shows, W would have to be .2 or less: i.e., E_2 would have to be weighted 4 times as strongly as E_1 instead of E_1 being weighted 1.5 times more than E_2. (Alternatively, W would have to be greater than 1.) This magnitude of error must be regarded as highly unlikely. An even more compelling point, however, is to ask what alternative to pooling would in fact be

UNCLASSIFIED

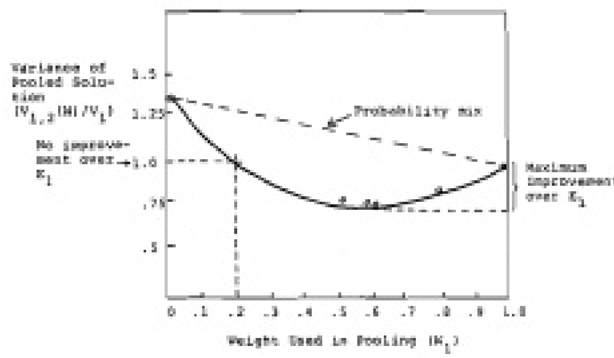


Figure 8. Ratio of the Pooled Variance ($V_{1,2}(W)$) to the Variance of the Most Precise Estimate Being Pooled (V_1) as a Function of Pooling Weight, When V_2/V_1=1.3, ρ=.34. Data Points from Sample 1 (from left to right): KAST, Unit Weighting, Variance-only Weights, Default Weights, Subjective Weights, MATE.

adopted. If the CO or a member of his staff were seriously considering values of W less than .2, the likely alternative to pooling for them would not be reliance on E_1 - but reliance on E_2. And in this case, pooling would be better no matter what positive value of W was used.

It is perhaps more realistic to suppose that the person bent on selecting a single estimate, rather than pooling, will be unsure which of two (or more) estimates to prefer. Under conditions where E_1 is on average better (i.e., $V_1 < V_2$), it is likely that he will sometimes select E_1 and sometimes E_2. The variance of this person's estimate will be a weighted average of the variances of E_1 and E_2 with weights determined by the probability (q) of selecting E_1:

$$\begin{aligned} V_{1,2}(q) &= qV_1 + (1-q)V_2 \\ V_{1,2}(q)/V_1 &= q + (1-q)p. \end{aligned} \tag{7}$$

(This formulation assumes that E_1 and E_2 are both unbiased, or equally biased, estimates of R; otherwise, the discrepancy between the means of E_1 and E_2 will add to the variance.) Thus, the variance for this probabilistically mixed outcome will fall somewhere on (or above) the straight dashed line in Figure 8 linking V_1 and V_2. Pooling, even with nonoptimal weights, will usually do better than E_1 (the better of the two estimates), but the best that unique selection can achieve is to equal E_1. Moreover, to the extent that a user would sometimes select the less good alternative, the amount of error in pooling weights that is consistent with the superiority of pooling increases.

Figure 9 shows, more generally, how the leeway for pooling superiority depends on the optimal pooling weight and on the accuracy of unique selection. From formulae (6) and (7), we have for any values of p and ρ:

$$V_{1,2}(W) \geq V_{1,2}(q) \quad \text{if and only if}$$

$$W \geq w^* + \sqrt{q(1-w^*)^2 + (1-q)w^{*2}} \quad \text{or} \tag{8}$$

$$W \leq w^* - \sqrt{q(1-w^*)^2 + (1-q)w^{*2}},$$

VALIDATION OF A TARGET RANGE DECISION AID
Marvin S. Cohen

UNCLASSIFIED

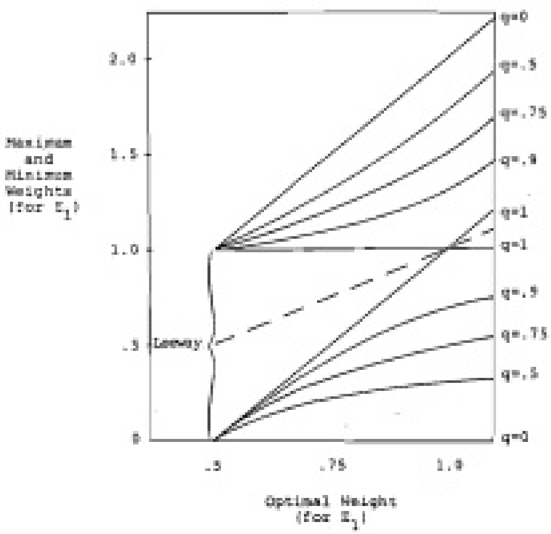


Figure 9. Maximum and Minimum E_1 Weights Required for Superiority of Pooling Over Unique Selection (Resulting in Probabilistic Mixing of E_1 and E_2), as a Function of the Optimal Weight and the Probability of Selecting E_1 (q)

where w* is the optimal weight as specified in formula (4). Suppose, for example, that the optimal pooling weight is .75 and that unique selection under those conditions means selecting E_1 75% of the time and E_2 25% of the time. Then the weight actually used in pooling could be as low as .32 or as high as 1.18 - and pooling would still be better than the strategy of unique selection. (When unique selection is perfect (q=1), the upper and lower bounds would be 1 and .5.) If the "true" weight were .9 and unique selection involved mistakenly picking E_2 only 10% of the time, the minimum pooling weight yielding an advantage for pooling remains quite low at .6. In short, even when there is significant uncertainty about pooling weights, and one estimate is significantly more accurate than the other, it is very unlikely that the strategy of selecting a single estimate will do better than pooling.

Note on Bias

In the discussion of sensitivity, we have dealt with the expected impact of pooling on random error, not on bias. Any pertinent information about systematic error in a specific ranging technique (e.g., Ekelund) can and should be used whether or not the technique is pooled with others. To the extent that bias in an estimate being pooled is not successfully corrected, the remaining bias is transmitted into the pooled solution in proportion to the weight associated with that estimate. But it will show up (in full) in range estimates that are not pooled.

ASSESSMENTS OF RANGE UNCERTAINTY

Perhaps as important as increasing the accuracy of range solutions is knowing just how accurate the solution is, at a given time. The pooling aid algorithm produces an estimate of error variance for the pooled solution. From this, intervals can be calculated which are believed to contain the true range with various degrees of certainty.

Table 5 enables us to assess the calibration of the range pooling aid with default parameters for 95% and 90% intervals of uncertainty. The percentages within box A should hover around 95%, corresponding to the proportion of the time the true range would be contained in a valid 95% interval. In box B, the percentages should approximate 90%. In both cases, calibration is not far off, but the predicted intervals are somewhat too narrow. Across all pooling combinations and samples, the true range falls inside the 95% interval 92% of the time, and within the 90% interval 89% of the time.

UNCLASSIFIED

	A		B	
	95% Interval		90% Interval	
Solutions Pooled	First Sample	Second Sample	First Sample	Second Sample
MATE KAST	90	92	87	88
MATE EKELUND	91	89	89	87
KAST EKELUND	93	96	91	95
MATE-KAST EKELUND	91	94	88	90
Overall	92		89	

Table 5. Proportion of Time True Range Falls Within Predicted Interval of Uncertainty

Deviations from normality and dependence of error (V_i) on range would be expected to distort pooling aid calibration. An additional factor, not reflected in Table 5, is the artificiality of exercise conditions in comparison to combat: The number of targets, their bearings, possible maneuvers, and maximum distance are all known in advance. This extra uncertainty should

VALIDATION OF A TARGET RANGE DECISION AID
Marvin S. Cohen

not affect improvements in accuracy due to pooling relative to use of simple estimates, but it affects the precision with which accuracy (of specific estimates and of the pooled solution) can be known. Techniques incorporating higher-order judgments (uncertainty about uncertainty) and which use the dispersion of range estimates as an additional clue to the spread of the pooled distribution are being applied to this problem (cf., Lindley, 1981; Winkler, 1981).

CONCLUSIONS

The pooling algorithm provides a systematic (yet quite simple) method for combining all the available information about target range: diverse range estimates, objective data on the accuracy and interrelations of the techniques, and judgments by the CO or other appropriate personnel concerning any of the above. Our preliminary quantitative validation confirms the efficacy of the approach: pooling improved range accuracy whether with objective inputs, subjective inputs, or a mix of the two. Moreover, the degree of accuracy achieved corresponded to the amount of pooling involved, and thus to the amount of information that was integrated. Accuracy increased in the following order:

- MATE, KAST, Ekelund, and the Command staff best guess
- Pooling MATE, KAST, and Ekelund with subjective weights
- Pooling MATE, KAST, and Ekelund with default weights
- Pooling the default pooled solution with the Command staff best guess

Sensitivity analysis suggests that, at any of these levels of pooling, significant errors in parameter estimates can occur without jeopardizing the advantage of pooling over what is being pooled. The conclusion, we should stress, is not that objective data are "better" than subjective judgment, or vice versa. Rather, it is that a prescriptive model which systematically constrains the form of the inference process while accommodating both objective data and individual judgments may be best of all.

Current work is focusing on the actual expected contribution of pooling to combat effectiveness. The time-course of uncertainty reduction about target range with and without the pooling aid is being analyzed within approach and attack exercises. Implications of improvements for the CO's tactical flexibility (e.g., earlier time of fire) and for the accuracy of assessments involved in tactical decisions will be examined. A prototype range-pooling aid, being implemented at the Naval Underwater Systems Center (Newport, R.I.), will be used for hands-on demonstrations with potential users. A critical topic for investigation will be the feasibility of eliciting the subjective judgments which have thus far only been simulated.

REFERENCES

Cohen, M.S. Decision support for attack submarine commanders: Target range pooling and attack planning (Technical Report 82-1). Falls Church, VA: Decision Science Consortium, Inc., April 1982.

Cohen, M.S., and Brown, R.V. Decision support for attack submarine commanders (Technical Report 80-11). Falls Church, VA: Decision Science Consortium, Inc., October 1980.

Cohen, M.S., and Brown, R.V. Application of decision analysis to ASW ranging and time of fire (U). Proceedings of the April 1981 Steering Group Meeting for the Advanced Attack Center (AAC) (U). Newport, RI: Naval Underwater Systems Center, 10 August 1981. Confidential

Cohen, M.S., Bromage, R.C., Chinnis, J.O., Jr., Payne, J.W., and Ulvila, J.W. A personalized and prescriptive decision aid (Technical Report 82-4). Falls Church, VA: Decision Science Consortium, Inc., July 1982. (a)

Cohen, M.S., Brown, R.V., Seaver, D.A., Ulvila, J.W., and Stillwell, W.G. Operability in attack submarine combat systems: An exploratory review (Technical Report 82-2). Falls Church, VA: Decision Science Consortium, Inc., April 1982. (b)

Dawes, R.M. The mind, the model, and the task. In Restle, F. et al., (Eds.) Cognitive theory, (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum, Asso., 1975, 119-130.

Dawes, R.M. The robust beauty of improper linear models in decision making. American Psychologist, July 1979, 34(7), 571-582.

Lindley, D.V. Reconciling continuous probability assessments (Technical Report 81-5). Falls Church, VA: Decision Science Consortium, Inc., October 1981.

Lindley, D.V., Tversky, A., and Brown, R.V. On the reconciliation of probability assessments. The Journal of the Royal Statistical Society, Series A, 1979, 142(2), 146-180.

Morris, P.A. Combining expert judgments: A Bayesian approach. Management Science, 1977, 23(7), 679-693.

Patterson, J.F., Randall, L.S., and Stewart, R.R. Advisory decision aids: A prototype. McLean, VA: Decisions and Designs, Inc., February 1981.

Winkler, R.L. Combining probability distributions from dependent information sources. Management Science, 1981, 27(4), 479-488.

NOTE

1. This work is part of a more general project, now in its third year, to develop computer-based decision support systems for attack submarine commanders (Cohen, 1982; Cohen and Brown, 1981, 1980). The work is sponsored by the Engineering Psychology Program, Office of Naval Research (ONR), in collaboration with the Naval Underwater Systems Center (NUSC, Newport, R.I.). A prototype range pooling aid is currently being implemented at a testbed facility at NUSC.

Appreciation is due to R.V. Brown, T. Downie, A.N.S. Freeling, H. Kettel, and L. Merchant-Gauder for their contributions to the research and data analysis. Constructive guidance has been provided by Dr. Martin Tolcott and Mr. Gerald Malecki of the Office of Naval Research. Appreciation is also due to Mr. David Barry, Dr. Albert Colella, Mr. Craig Gardner, CAPT

VALIDATION OF A TARGET RANGE DECISION AID

Marvin S. Cohen

John Gilchrist (U.S. Navy-Ret.), Mr. Wayne King, CDR Richard Pariseau (U.S. Navy-Ret.), Mr. Francis Spicola, and CAPT Leonard Stoehr (U.S. Navy-Ret.)

DESCRIPTORS: Decision Aids, Antisubmarine Warfare, Range Estimation, Decision Aid Validation, Rangex Data, Subjective Decision Aid Inputs

COMMENTARY

Richard R. Pariseau
System Control Technology

The paper presents a useful approach to a very real problem. The reality of the problem is appreciated by the fact that during submarine vs. submarine exercises the average elapsed time from initial detection until weapon launch is approximately two hours. In passive target tracking, target range continues to be the most uncertain value and largely responsible for the time interval. When torpedoes are fired the target damage ratio has historically been low. For various reasons the number of successful torpedo firings today is only about fifty percent. It is obvious therefore that the problem is relevant and decision aiding is necessary.

A major problem facing technicians who are developing decision aids is the presentation of the results of their effort. The decision maker, in this case a submarine commanding officer, must feel comfortable with the data which he is presented. It is therefore necessary to the decision analyst to be concerned with FORMAT, VALIDATION and the INDIVIDUALITY of the decision makers.

The data must be presented to the decision maker in a format which is familiar, easily understood, quickly comprehended and unambiguous.

When the accuracy of the input data from which decisions must be made is of unknown accuracy it is natural and necessary for the decision maker to spend a significant amount of time trying to validate ie., determine the accuracy of, the input data. A less satisfying alternative but complementarily addition to validation is sensitivity of the results to input data errors. It must always be remembered that a perfect solution is not necessarily the goal of the decision maker, but a comfortable awareness of the validity of the data upon which a decision is based is always important.

The subjectivity of the individual decision makers initially impacts acceptance of the decision aid and ultimately the frequency of its usage and reliance upon the output of the aid once it is installed aboard the ships. There are several ways of handling the subjectivity requirements from basic capability for data call-up at any level upon command to personal cassettes carries aboard by individual commanding officers which present and validate information in a personal manner. A single solution will not be optimal for all decision aids, however each decision aid must be developed in consideration of the subjectivity of the decision maker.

UNCLASSIFIED