ESTIMATION OF AIM-ERROR CORRELATION, AIMING DISPERSION, AND BALLISTIC DISPERSION

Joseph Bram

Research Contribution 202

Center for Naval Analyses

Systems Evaluation Group

an affiliate of the University of Rochester

1401 Wilson Boulevard, Arlington, Virginia 22209

None
Security Classification

DOCUMENT CONTROL DATA - R & D

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

1. ORIGINATING ACTIVITY (*Corporate author*)
Systems Evaluation Group, Center for Naval Analyses
an affiliate of the University of Rochester

2a. REPORT SECURITY CLASSIFICATION
None

2b. GROUP
None

3. REPORT TITLE
Estimation of Aim-Error Correlation, Aiming Dispersion, and Ballistic Dispersion

4. DESCRIPTIVE NOTES (*Type of report and inclusive dates*)
Research Contribution - December 1971

5. AUTHOR(S) (*First name, middle initial, last name*)
Bram, Joseph

6. REPORT DATE
December 1971

7a. TOTAL NO. OF PAGES
30

7b. NO. OF REFS
3

8a. CONTRACT OR GRANT NO.
N00014-68-A-0091

b. PROJECT NO.
---

c.
---

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
Systems Evaluation Group
Research Contribution 202

9b. OTHER REPORT NO(S) (*Any other numbers that may be assigned this report*)
None

10. DISTRIBUTION STATEMENT
Approved for Public Release; Distribution Unlimited.

11. SUPPLEMENTARY NOTES
None

12. SPONSORING MILITARY ACTIVITY
Office of Naval Research
Department of the Navy
Washington, D.C. 20350

13. ABSTRACT

Assuming that the x-coordinates of impact points of rounds fired at a target centered at the origin, $x = o$, are the sums of correlated aim errors and independent ballistic errors, we show how the likelihood function of the unknown statistical parameters, associated with a sample $x_1, x_2, \ldots, x_N$ of impact points from a single burst, can be quickly and simply computed. The likelihood function is used here as the basis for estimating the parameters, and the results of our estimates, applied to a large number of computer simulations, are summarized.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6801

None
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
correlated aim errors independent ballistic errors aiming dispersion ballistic dispersion						

Center for Naval Analyses

1401 Wilson Boulevard Arlington Virginia 22209 703/524 9400

(SEG) 278-72
29 Feb 1972

From: Director, Systems Evaluation Group
To: Distribution List

Subj: Center for Naval Analyses Research Contribution 202; forwarding of

Encl: (1) CRC 202, "Estimation of Aim-Error Correlation, Aiming Dispersion, and Ballistic Dispersion," by J. Bram, SEG, December 1971

1. Enclosure (1) is forwarded as a matter of possible interest.

2. Research Contributions are distributed for their potential value in other studies and analyses. They do no necessarily represent the opinion of the Department of the Navy.

3. The enclosure has been approved for public release.

Phil E. DePOY

PHIL E. DePOY
Director

Distribution List:
Reverse page

(SEG)278-72
29 Feb 1972

Subj: Center for Naval Analyses Research Contribution 202; forwarding of

---

DISTRIBUTION LIST

Department of the Navy:

Ass't SecNav (R&D)
ONR
OpNav: Op-09BH Op-093 Op-980 Op-982 Op-32 Op-55
Op-96 Op-098 Op-981 Op-03 Op-05 ,
SNDL Part 1:

CINCPACFLT	COMOPTEVFOR
CINCLANTFLT	DEPCOMOPTEVFORPAC
CINCUSNAVEUR	AIRTEVRON 4 (VX-4)
COMFIRSTFLT	AIRTEVRON 5 (VX-5)
COMSECONDFLT	COMCRUDESLANT
COMSIXTHFLT	COMCRUDESPAC
COMSEVENTHFLT	

SNDL Part 2:

NRL	NAVAIRTESTCEN
USNA ANNA	NAVWPNSYSANALO
NAVPGSCOL	NAVAVIONICFAC
NAVWARCOL	NAVSHIPMISYSENGSTA
FAAWTRACENLANT	CHNAVMAT
FAAWTRACENPAC	NAVAIRSYSCOMHQ
NAVAIRDEVCEN	NAVORDSYSCOMHQ
NAVWPNSCEN	NAVSHIPSYSCOMHQ
NWL	
NOL	

Other Department of Defense Agencies:

DDR&E	Dept of Army (Attn: Adj Gen'l)(6)
SENAVMBR, WSEG	Dept of Air Force (Attn: PDO 4008)(9)
DIR, ARPA	Superintendent, Air University
ADMIN, DDC (12)	Air Proving Ground Center, Eglin AFB

Other Agencies:

ANSER	RAND
APL/JHU	SRI
IDA	Cornell Aeronautical Lab., Inc.
ORL	Battelle Memorial Institute
RAC	University of Rochester (3)

CENTER FOR NAVAL ANALYSES

Systems Evaluation Group

Research Contribution 202

ESTIMATION OF AIM-ERROR CORRELATION, AIMING DISPERSION, AND BALLISTIC DISPERSION

December 1971

Joseph Bram

Work conducted under contract N00014-68-A-0091

Enclosure (1) to (SEG)278-72 dated 29 February 1972

ABSTRACT

Assuming that the x-coordinates of impact points of rounds fired at a target centered at the origin, $x = 0$, are the sums of correlated aim errors and independent ballistic errors, we show how the likelihood function of the unknown statistical parameters, associated with a sample $x_1, x_2, \ldots, x_N$ of impact points from a single burst, can be quickly and simply computed. The likelihood function is used here as the basis for estimating the parameters, and the results of our estimates, applied to a large number of computer simulations, are summarized.

(REVERSE BLANK)

TABLE OF CONTENTS

(REVERSE BLANK)

ESTIMATION OF AIM-ERROR CORRELATION, AIMING DISPERSION, AND BALLISTIC DISPERSION

INTRODUCTION

The importance of obtaining efficient estimates of the statistical parameters describing a rapid-fire gun system is well recognized. Such estimates are necessary in research studies of future systems and tactics, and also for the design of closed-loop "filters" to effect corrections at the gun site in actual firings, (references (b) and (c)). In reference (a) we began the study of the estimation problem by investigating what could be achieved from the traditional sample covariances of a time series $x_1, x_2, \ldots x_N$, and we showed that for high correlations ($\rho \approx 1$) in the aim error, the aiming dispersion σ_A^2 was virtually impossible to estimate from a single burst. Since rapid-fire guns have a high ρ, the results of reference (a) demonstrate the need for numerous independent experiments to estimate the parameters for any such gun system.

However, the information provided in reference (a) does not necessarily apply to cases with moderate ρ, or moderate burst lengths, and no attempt was made in reference (a) to provide formulas for estimating ρ and σ_A^2. In this paper, we show how to calculate the likelihood function associated with a sample $x_1, x_2, \ldots, x_N$. This function is then used to provide Bayes estimates of the parameters, as well as their second moments. The latter quantities permit us to obtain approximate confidence intervals for the parameters.

(Familiarity with reference (a) is not required in reading this paper.)

THE MODEL AND THE PROBLEM

With a target centered at the origin of a 1-dimensional coordinate system, we imagine a burst of N rounds impacting in points $x_1, x_2, \ldots, x_N$. (For a 2-dimensional picture, the x's may denote the azimuth errors and equally well, the elevation errors, treated separately and independently.) The model consists of the assumptions that each x_n is the sum,

$$x_n = \alpha_n + \beta_n \tag{1}$$

of two stationary Gaussian processes $\{\alpha_n\}$, $\{\beta_n\}$, independent of each other, in which $\{\alpha_n\}$ represents the aim error, means zero, and covariances

$$E[\alpha_n \alpha_{n+k}] = \sigma_A^2 \rho^{|k|} , \tag{2}$$

and $\{\beta_n\}$ represents the ballistic error, means zero, and covariances

$$E[\beta_n \beta_{n+k}] = \sigma_B^2 \delta_{o,k} . \tag{3}$$

The process $\{\beta_n\}$ is often called "white noise." There are exactly three parameters: σ_A^2, the aiming dispersion; σ_B^2, the ballistic dispersion; and ρ, the round-to-round aim correlation.

The problem is to estimate the values of the 3 parameters, and to obtain confidence intervals for them from impact data. Equivalently, from a Bayes viewpoint, we want the posterior probability density in the parameter space, given a single sample $x_1, x_2, \ldots, x_N$ of impact points from a burst of N rounds.

According to this model, the gun is unbiased, $E(x_n) = 0$ for each n. The possibility of a non-zero bias is of no importance in the design of a gun filter, because any such possible bias will be removed by the filter, (reference (b)). Furthermore, in sampling experiments on the firing range, the data $x_1, \ldots, x_N$ can be replaced by $x_1 - \bar{x}$, $x_2 - \bar{x}$, $\ldots, x_N - \bar{x}$, where $\bar{x}$ is the arithmetic mean. The only parameters of importance are σ_A^2, σ_B^2, and ρ.

We confine ourselves to the estimation problem based on a single burst $x_1, \ldots, x_N$, since there is no problem in combining the results of independent experiments. If $L_\nu(\sigma_A^2, \sigma_B^2, \rho)$ denotes the likelihood function of the νth independent trial (with the same σ_A^2, σ_B^2, ρ) for $\nu = 1, 2, \ldots, r$, then, of course, the combined likelihood function is simply the product

$$L\left(\sigma_A^2, \sigma_B^2, \rho\right) = \prod_{\nu=1}^{r} L_\nu\left(\sigma_A^2, \sigma_B^2, \rho\right) \quad .$$

THE LIKELIHOOD FUNCTION

From the last section, and equations (1), (2), (3), we see that the sample vector $(x_1, x_2, \ldots, x_N)$ is Gaussian, with mean $(0, 0, \ldots, 0)$, and covariances

$$\operatorname{cov}(x_i, x_j) = \sigma_A^2 \rho^{|i-j|} + \sigma_B^2 \delta_{ij} \tag{4}$$

If $C = \{c_{ij}\}$ is the $N \times N$ covariance matrix, $c_{ij} = \operatorname{cov}(x_i, x_j)$, then the density function for $(x_1, \ldots, x_N)$ is

$$\tau\left(x_1, \ldots, x_N | \sigma_A^2, \sigma_B^2, \rho\right) = \frac{\left[\det C\right]^{-\frac{1}{2}}}{(2\pi)^{N/2}} \exp\left[\frac{-x^t C^{-1} x}{2}\right] \tag{5}$$

where x is the $N \times 1$ matrix with elements $x_1, \ldots, x_N$; x^t is the transpose of x; C^{-1} is the matrix inverse to C; and det C is the determinant of C. The likelihood function is precisely the quantity in equation (5),

$$L(\sigma_A^2, \sigma_B^2, \rho) = \wp(x_1, \ldots, x_N | \sigma_A^2, \sigma_B^2, \rho) \tag{6}$$

(or any constant multiple thereof), regarded as a function of the (Greek-letter) parameters. The observations $x_1, \ldots, x_N$ are fixed, of course, for any one experiment.

But the form of $\wp$ in equation (5) is practically impossible to evaluate, even on a high-speed computer. The matrix C is easy to write, equation (4), but C^{-1} cannot be written in terms of the parameters, and inverting large matrices on a computer is time-consuming and subject to gross round-off errors. Therefore, we again consider the probability density of $(x_1, \ldots, x_N)$ and write, suppressing the Greek letters,

$$\wp(x_1, \ldots, x_N) = \wp_1(x_1)\, \wp_2(x_2|x_1)\, \wp_3(x_3|x_1, x_2) \ldots \wp_N(x_N|x_1, \ldots, x_{N-1}) \tag{7}$$

where each $\wp_j$ is the conditional density of x_j given $x_1, \ldots, x_{j-1}$. Each $\wp_j$ is a Gaussian density, and we therefore need only the two quantities (for each $j = 1, 2, \ldots, N$),

$$\hat{x}_j = E\left[x_j | x_1, \ldots, x_{j-1}\right] \tag{8}$$

and

$$\sigma_j^2 = E\left[(x_j - \hat{x}_j)^2 | x_1, \ldots, x_j\right], \tag{9}$$

and we can then write

$$\wp_j(x_j | x_1, \ldots, x_j) = \frac{1}{\sqrt{2\pi}\,\sigma_j} \exp\left[-\frac{1}{2\sigma_j^2}(x_j - \hat{x}_j)^2\right] \tag{10}$$

From equation (6), (7), (10), we can write

$$L(\sigma_A^2, \sigma_B^2, \rho) = \prod_{j=1}^{N} \frac{1}{\sigma_j} \exp\left[-\frac{1}{2\sigma_j^2}(x_j - \hat{x}_j)^2\right] \tag{11}$$

where we have discarded the constant multiples $(2\pi)^{-1/2}$, and in which the $\hat{x}_j$ depend on $x_1, \ldots, x_{j-1}$, and $\sigma_A^2, \sigma_B^2, \rho$, whereas the σ_j depend only on $\sigma_A^2, \sigma_B^2, \rho$!

(The proof of the last clause can be carried out along the following lines. The conditional density of x_j given $x_1, \ldots, x_{j-1}$ can be written

$$\varphi_j(x_j | x_1, \ldots, x_{j-1}) = k(x_1, \ldots, x_{j-1})\, \psi(x_1, \ldots, x_j)$$

where ψ denotes the density of $(x_1, \ldots, x_j)$, and $k(x_1, \ldots, x_{j-1})$ is the normalizing factor obtainable from $\int \varphi_j\, dx_j = 1$. Now ψ has the form given by the r.h.s. of equation (5), with N replaced by j. The x_j appears only in the exponent, so we can write

$$\varphi_j(x_j | x_1, \ldots, x_{j-1}) = k'(x_1, \ldots, x_j) \cdot e^{-1/2\, Q}$$

where

$$Q = (x_1, \ldots, x_j)\, A \begin{bmatrix} x_1 \\ \vdots \\ x_j \end{bmatrix}$$

and A is the matrix inverse of the j-rowed covariance matrix of $x_1, \ldots, x_j$. By familiar methods, Q can be written

$$Q = Q_o(x_1, \ldots, x_{j-1}) + A_{jj}\, (x_j - \hat{x}_j)^2$$

so that

$$\sigma_j^2 = A_{jj}^{-1} ,$$

where A_{jj} is the (j, j) element of A. For our purposes, the main point is that σ_j^2 does not depend on $x_1, \ldots, x_{j-1}$.)

The interesting consequence of the last remark is that

$$E\left[(x_j - x_j)^2\right] = \sigma_j^2 , \tag{12}$$

so that σ_j^2 is the (unconditional) mean square difference between x_j and $\hat{x}_j(x_1, \ldots, x_{j-1})$. Now the problem of calculating $\hat{x}_j$ and σ_j^2 in our context is a special case of the Kalman filter theory, which is treated, for example, in reference (c), and which we summarize in appendix A, for completeness.

Calculating the Likelihood Function

For a fixed sample $x_1, x_2, \ldots, x_N$ and arbitrary σ_A, σ_B, ρ, we require L, given in equation (11), or equivalently,

$$h(\sigma_A, \sigma_B, \rho) = \log_e L \ . \tag{13}$$

Appendix A shows how the σ_j and $\hat{x}_j$ of equation (11) can be calculated recursively. The essential equations are (A-17), (A-18), (A-20), (A-21), and (A-22), which we repeat here.

$$\left.\begin{aligned} \hat{x}_1 &= 0 \\ u_o &= 0 \end{aligned}\right\} \tag{14}$$

$$\sigma_n^2 = (\sigma_A^2 + \sigma_B^2 - u_{n-1}) \tag{15}$$

$$\gamma_n' = \frac{\rho\ (\sigma_A^2 - u_{n-1})}{\sigma_n^2} \tag{16}$$

$$u_n = \rho\ (\rho - \gamma_n')\ u_{n-1} + \rho\ \sigma_A^2\ \gamma_n' \tag{17}$$

$$\hat{x}_{n+1} = (\rho - \gamma_n')\ \hat{x}_n + \gamma_n'\ x_n \tag{18}$$

In figure 1, we have displayed a flow diagram for computing $h(\sigma_A, \sigma_B, \rho)$. The letters V, H, C, U, Y stand for σ_n^2, h, γ_n', u_n, $\hat{x}_n$. This efficient algorithm for h can be made faster by not calculating $\log(V) = \log(\sigma_n^2)$ for each index, but rather by forming the products $\prod_{n=1}^{j} \sigma_n^2$ recursively (guard against overflow!), and then calculating $\log\left(\prod_{1}^{N} \sigma_n^2\right)$ just once, for each $(\sigma_A, \sigma_B, \rho)$.

Estimating the Parameters

For each parameter point $(\sigma_A, \sigma_B, \rho)$ — a point in 3-space — we can evaluate h and therefore $L = e^h$, from equation (13). But in place of ρ, which we expect to be in the interval $0 < \rho < 1$, we prefer to use the variable ζ, where

$$\zeta = \frac{1}{2} \log \frac{1+\rho}{1-\rho} \ . \tag{19}$$

This is equivalent to

$$\rho = \frac{e^{2\zeta} - 1}{e^{2\zeta} + 1} = \tanh(\zeta) \tag{20}$$

and we take our parameter space to be the set of all points $(\sigma_A, \sigma_B, \zeta)$ in R_3 for which all three coordinates are ≥ 0 .

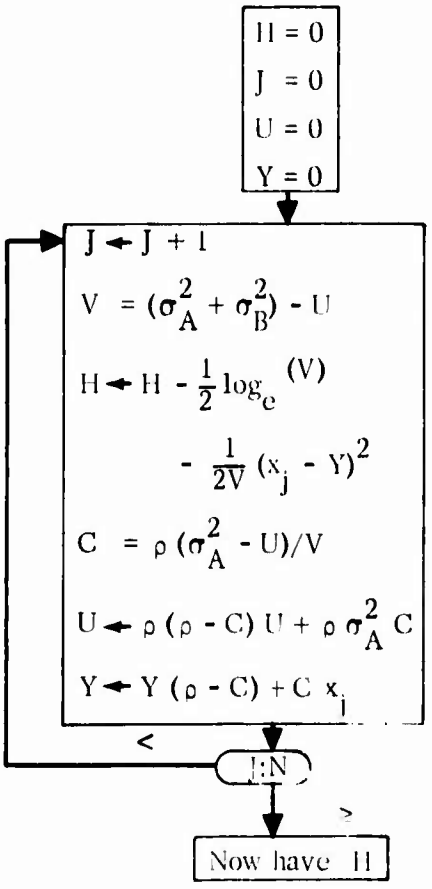

FIG. 1: ROUTINE FOR CALCULATING $h = \log_e L$

If $\varphi\ (\sigma_A, \sigma_B, \zeta)$ denotes our prior probability density, then after the sample $(x_1, \ldots, x_N)$ is observed, our posterior density is

$$\hat{\varphi}\ (\sigma_A, \sigma_B, \zeta) = C\ \varphi\ L = C\ \varphi\ e^h \tag{21}$$

where C is the constant required to make equation (22) true:

$$\iiint C\ \varphi\ e^h\ d\sigma_A\ d\sigma_B\ d\zeta = 1\ , \tag{22}$$

i.e.,

$$C = \frac{1}{\iiint \varphi\ e^h\ d\sigma_A\ d\sigma_B\ d\zeta}\ .$$

We take for our estimates:

$$\hat{\sigma}_A = E(\sigma_A) = \iiint \sigma_A\ \hat{\varphi}\ d\sigma_A\ d\sigma_B\ d\zeta\ ,$$

$$\hat{\sigma}_B = E(\sigma_B) = \iiint \sigma_B\ \hat{\varphi}\ d\sigma_A\ d\sigma_B\ d\zeta\ ,$$

$$\hat{\zeta} = E\ (\zeta) = \iiint \zeta\ \hat{\varphi}\ d\sigma_A\ d\sigma_B\ d\zeta,$$

and then, from equation (20),

$$\hat{\rho} \underset{\text{def}}{=} \tanh\ (\hat{\zeta})\ .$$

We have used for φ, the constant (uniform) density for $0 \le \sigma_A \le D, 0 \le \sigma_B \le D$, $0 < \zeta \le 3$, where D^2 is four times the 0th sample covariance, i.e., we put

$$D = 2 \left[\frac{1}{N} \sum_{i=1}^{N} x_i^2 \right]^{\frac{1}{2}}\ . \tag{23}$$

We note that $E\left[\frac{1}{N}\sum_{1}^{N} x_i^2\right] = \sigma_A^2 + \sigma_B^2$, and our D will be on the order of $2(\sigma_A^2 + \sigma_B^2)^{1/2}$.

The interval $0 \le \zeta < 3.0$ corresponds to $0 < \rho \le 0.995^{+}$.

With these assumptions on our prior, we are in a position to consider the evaluation of integrals of the form

$$\int\int\int F(\sigma_A, \sigma_B, \zeta)\, \hat{\varphi}(\sigma_A\, \sigma_B\, \zeta)\, d\sigma_A\, d\sigma_B\, d\zeta \quad .$$

We are principally interested in 9 such F's, viz., $F_1 = \sigma_A$; $F_2 = \sigma_B$; $F_3 = \zeta$; $F_4 = \sigma_A^2$; $F_5 = \sigma_A\, \sigma_B$; etc., i.e., we want the 1st and 2nd moments of σ_A, σ_B, ζ, and more important, the 2nd moments about the means. For example, with $F = \sigma_A\, \zeta$, our integral is, using equation (21):

$$\frac{\int_0^3\int_0^D\int_0^D \sigma_A\, \zeta\, e^{h(\sigma_A, \sigma_B, \zeta)}\, d\sigma_A\, d\sigma_B\, d\zeta}{\int_0^3\int_0^D\int_0^D 1\, e^{h(\sigma_A, \sigma_B, \zeta)}\, d\sigma_A\, d\sigma_B\, d\zeta} \quad .$$

To calculate all the required integrals, of which there are now 10, with the 10th F being $F = 1$, we used the Monte Carlo procedure, choosing 500 points $(\sigma_A, \sigma_B, \zeta)$ uniformly randomly in the region $0 \le \sigma_A \le D$, $0 \le \sigma_B \le D$, $0 \le \zeta < 3$, and we used the approximation

$$\int\int\int F_j\, \hat{\varphi}\, d\sigma_A\, d\sigma_B\, d\zeta = \frac{\sum_{\nu=1}^{500} F_j(\omega_\nu)\, e^{h_\nu}}{\sum_{\nu=1}^{500} 1\, e^{h_\nu}}$$

for $j = 1, 2, \ldots, 9$, where ω_ν denotes the νth random point $(\sigma_A, \sigma_B, \zeta)$. (Since the h_ν were on the order of $h_\nu = -200$, the sums in the previous equation were replaced by sums of the form

$$\sum_\nu F(\omega_\nu) \exp(h_\nu - h^*),$$

where $h^* = \max_\nu h_\nu$, and, of course, the constant factor $\exp(-h^*)$ cancels in numerator and denominator.)

EXPERIMENTAL PROCEDURE

We chose 16 parameter combinations, viz., $\sigma_A = 6, 4, 2, 1$, with $\rho = .98, .90, .80, .50$ ($4\sigma_A$'s times 4 ρ's), with $\sigma_B = 2$ throughout. For each parameter combination we took 5 independent samples $x_1, \ldots, x_N$ with $N = 100$ in every case. So we collected 80 independent samples in all. Each sample was generated by the computer, programmed to simulate the process $\{x_n\}$ according to the model described earlier.

The computer calculated and printed out the estimates $E(\sigma_A)$, $E(\sigma_B)$, $E(\zeta)$ (expectations with respect to our posterior density in the parameter space), together with the second moments about the means, as described in the last section. This was done for each of the 80 independent samples.

Experimental Results

The means $\hat{\sigma}_A$, $\hat{\sigma}_B$, $\hat{\zeta}$ (and also $\hat{\rho} = \tanh(\hat{\zeta})$), and the standard deviations $\sqrt{D^2(\sigma_A)}, \sqrt{D^2(\sigma_B)}, \sqrt{D^2(\zeta)}$ that resulted from the 80 samples are all presented in appendix B. The mixed moments such as $E(\sigma_A \sigma_B)$ and the covariances were of less importance and are not given here. (The covariances indicated, as expected, that the estimates $\hat{\sigma}_A$, $\hat{\sigma}_B$, $\hat{\zeta}$ were not independent of each other.)

The main purpose of computing the second central moments was to provide a method for calculating confidence intervals for the parameters σ_A, σ_B, ζ. Let σ_A denote the true value, $\hat{\sigma}_A$ its estimate, and $D^2(\sigma_A)$ the second central moment (with respect to the posterior density associated with a sample $x_1, \ldots, x_N$).

Define

$$x_A = \frac{(\hat{\sigma}_A - \sigma_A)^2}{D^2(\sigma_A)} \qquad (24)$$

We have 80 independent samples of this quantity when we aggregate all the σ_A's and ρ's. We calculated these quantities - or rather their square roots - and grouped them in intervals, and compared the frequencies with those obtainable on the hypothesis that x_A is distributed like chi-square with 1 degree of freedom. The comparison indicated that the latter hypothesis is a good assumption, i.e., that $\sqrt{x_A}$ is distributed like the absolute value of the standard unit normal variate (so x_A is chi-square with 1 d.f.).

The table below shows the observed and expected frequencies.

$$\text{Frequencies of } \sqrt{x_A} = \frac{|\hat{\sigma}_A - \sigma_A|}{\sqrt{D^2(\sigma_A)}}$$

Interval	Observed = O	Expected = E	O - E
0 - .5	28	30	-2
.5 - 1.0	21	25	-4
1.0 - 1.5	18	14	4
1.5 - 2.0	6	7	-1
> 2.0	7	4	3

The chi-square test applied here gives

$$Q = \sum \frac{(O-E)^2}{E} = 4.30 ,$$

which is small for four degrees of freedom, and indicates a satisfactory fit.

We did exactly the same sort of calculation for

$$x_B = \frac{(\hat{\sigma}_B - \sigma_B)^2}{D^2(\sigma_B)} \tag{25}$$

and for

$$x_\zeta = \frac{(\hat{\zeta} - \zeta)^2}{D^2(\zeta)} \tag{26}$$

and found excellent agreement with the hypothesis that x_B and x_ζ are also distributed like chi-square with one degree of freedom.

Using the division points 0, .5, 1.0, 1.5, 2.0, ∞, as before, the numbers of observations of $\sqrt{x_B}$ in the 5 intervals were 35, 27, 11, 3, 4, yielding O - E = 5, 2, -3, -4, 0. The chi-square Q turns out to be Q = 3.91 (four degrees of freedom). The corresponding numbers of observations of $\sqrt{x_\zeta}$ in the 5 intervals were 29, 25, 12, 9, 5, and the corresponding Q was 1.13 .

As a result, we are justified in saying that each of x_A, x_B, x_ζ is like chi-square with 1 d.f. Accordingly, if λ_p denotes p - percent value of a normal deviate, i.e.,

$$\frac{p}{100} = \frac{2}{\sqrt{2\pi}} \int_{\lambda_p}^{\infty} e^{-\frac{1}{2} t^2} \, dt \quad ,$$

then we can use the intervals

$$|\sigma_A - \hat{\sigma}_A| < \lambda_p \sqrt{D^2(\sigma_A)} \quad ,$$

$$|\sigma_B - \hat{\sigma}_B| < \lambda_p \sqrt{D^2(\sigma_B)} \quad ,$$

$$|\zeta - \hat{\zeta}| < \lambda_p \sqrt{D^2(\zeta)} \quad ,$$

as (100 - p) percent confidence intervals (though not all simultaneously).

As an example, take p = 10 percent. Then $\lambda_p = 1.64$. Suppose that we observe (calculate) $\hat{\zeta} = 1.26$, $\sqrt{D^2(\zeta)} = .57$. Then we can say

$$|\zeta - 1.26| < (1.64)(.57)$$

or

$$.32 < \zeta < 2.20$$

at the 90 percent level. Since $\rho = \tanh(\zeta)$ is a monotonic function of ζ, we can also say

$$.31 < \rho < .97$$

at the 90 percent level. If we follow such a procedure, then in a long series of such independent experiments and assertions, we will be wrong about 10 percent of the time.

CONCLUDING REMARKS

We have described a model for the miss distances (plus or minus) $x_1, \ldots, x_n$ arising from a single burst of a rapid-fire gun. This model contains 3 parameters: σ_A, σ_B, ρ. The problem is to estimate these 3 parameters from one or more samples $(x_1, \ldots, x_n)$, though we concentrated on the case of one sample. (There is no problem in combining independent samples, because the required likelihood function is obtainable from those of each single sample by multiplying the latter together.)

We recommended a Bayes type of procedure, in which we used a uniform prior density over a certain 3-dimensional interval in the parameter space of σ_A, σ_B, $\zeta = \tanh(\rho)$. The posterior density is then the product of the likelihood function and the prior, apart from a multiplicative constant factor. For our estimates, we took the expectation of σ_A and σ_B and ζ, relative to our posterior density. This procedure is precisely the optimal procedure when our loss function is the mean square error. More precisely, suppose we concentrate on σ_A, for example, and somebody proposes a rule $\psi(x_1, \ldots, x_n)$ for estimating σ_A, i.e., observe $(x_1, \ldots, x_n)$, form the real number $\psi(x_1, \ldots, x_n)$ and take this to be our estimate $\hat{\sigma}_A$ of σ_A.

For any fixed σ_A, σ_B, ζ, the totality of all possible estimates $\hat{\sigma}_A = \psi(x_1, \ldots, x_n)$ -- $(x_1, \ldots, x_n)$ varying according to its p.d.f.-- has a certain p.d.f. We form the differences-squared, $(\hat{\sigma}_A - \sigma_A)^2$, and their expected value:

$$\rho(\sigma_A, \sigma_B, \zeta; \psi) = \underbrace{\int \ldots \int}_{N} [\psi(x_1, \ldots, x_n) - \sigma_A]^2$$

$$\times f(x_1, \ldots, x_n | \sigma_A, \sigma_B, \zeta) dx_1, \ldots, dx_N$$

where f is the probability density of $(x_1, \ldots x_n)$ given σ_A, σ_B, ζ. We assume that this function ρ, depending on the parameters and on ψ, is the loss that we wish to be small. Let ψ_o denote the rule proposed in this paper. To say that ψ is better than ψ_o is not yet a meaningful statement, for this statement, according to our assumption, ought to mean that

$$\rho(\sigma_A, \sigma_B, \zeta; \psi) < \rho(\sigma_A, \sigma_B, \zeta; \psi_o) .$$

But for which values of $(\sigma_A, \sigma_B, \zeta)$? We will not know, in practice, what the parameters are--if we did, there would be no problem. One possible way to make precise the statement "ψ is better than ψ_o" is to use the definition: for <u>all</u> possible values of $(\sigma_A, \sigma_B, \zeta)$, we have

$$\rho(\sigma_A, \sigma_B, \zeta; \psi) \leq \rho(\sigma_A, \sigma_B, \zeta; \psi_o) ,$$

and for some $(\sigma_A, \sigma_B, \zeta)$ the inequality is strict.

But it is not hard to prove the following for our problem: let $\pi(\sigma_A, \sigma_B, \zeta)$ denote any prior probability density that is non-zero in every region of 3-space where $\sigma_A > 0$, $\sigma_B > 0$, and let ψ_o be the function for which

$$\iiint \rho(\sigma_A, \sigma_B, \zeta;\ \psi)\, \varphi\ (\sigma_A, \sigma_B, \zeta)\ d\sigma_A d\sigma_B d\zeta$$

is smallest. This rule ψ_o will turn out to be

$$\psi_o(x_1, \ldots, x_n) = \frac{\iiint \sigma_A\ \varphi\ L(\sigma_A, \sigma_B, \zeta)\ d\sigma_A d\sigma_B d\zeta}{\iiint \varphi\, L(\sigma_A \sigma_B \zeta)\ d\sigma_A d\sigma_B d\zeta} \quad .$$

Then there is no other ψ about which one can say, "ψ is better than ψ_o." In more technical language, ψ_o is "admissible."

In our situation, we took a prior $\varphi(\sigma_A, \sigma_B, \zeta)$ that was zero for $\zeta < 0$ or $\zeta > 3.00$ and for $\sigma_A > 2\left[\sum_1^N x_i^2\right]^{1/2} = D$ and for $\sigma_B > D$. The reason was that we knew that $\zeta > 0$, at least in our simulations and very likely in practice (although it may be safer to take the interval $-.5 < \zeta < 3.0$ in practice), and in no simulation, of which we have run about 200, have we ever observed a case with $\sum_1^N x_i^2 < \frac{1}{4}(\sigma_A^2 + \sigma_B^2)$. So the latter event is one with very small probability, and we ignore that possibility in our procedure.

It is, consequently, unlikely that anyone will produce a rule ψ that is (uniformly, for all σ_A, σ_B, ζ) better than our rule ψ_o, and of course, the same also applies to the other parameters σ_B and ζ.

REFERENCES

(a) Bram, J., CNA Research Contribution 183, "Estimating Accuracy Parameters of Rapid-Fire Guns," Unclassified 27 Apr 1971

(b) Bram, J., Memorandum for Director, SEG, (SEG)74-71, "Optimal Gunnery Correction When Bias is Present," Unclassified 26 Apr 1971

(c) Corn, R. J., Memorandum for Director, SEG, (SEG)37-71, "An Analysis of Closed-Loop Control of Gun Systems," Unclassified 2 Mar 1971

APPENDIX A

We consider the stationary Gaussian process $x_1, x_2, \ldots, x_n, \ldots$ with means zero, and covariances

$$E[x_i x_j] = \sigma_A^2 \rho^{|i-j|} + \sigma_B^2 \delta_{ij} \tag{A-1}$$

with σ_A^2, σ_B^2, ρ assumed known. The problem is to calculate recursively

$$\hat{x}_n = E[x_n | x_1, \ldots, x_{n-1}] \tag{A-2}$$

and

$$\sigma_n^2 = E[(x_n - \hat{x}_n)^2] \tag{A-3}$$

It is well known that for Gaussian variables $\{x_n\}$, the $\hat{x}_n$ of (A-2), is a linear combination of $x_1, \ldots, x_{n-1}$, and in fact, it is that linear combination for which

$$Q(\hat{x}_n) = (x_1, \ldots, x_{n-1}, \hat{x}_n) A_n \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ \hat{x}_n \end{pmatrix} \tag{A-4}$$

is a minimum. Here A_n is the matrix inverse to the covariance matrix of $x_1, \ldots, x_n$. It follows that

$$\hat{x}_n = -\sum_{j=1}^{n-1} \frac{A_{nj}}{A_{nn}} x_j \tag{A-5}$$

where $(A_{n1}, \ldots, A_{nn})$ is the nth row of A_n. But (A-5) is not very useful, since we don't have simple formulas for the A_{nj}.

Let $C_{ij} = \text{cov}(x_i, x_j)$, and form the sums

$$\sum_{j=1}^{n-1} C_{ij} A_{nj} = \delta_{in} - C_{in} A_{nn} . \tag{A-6}$$

Therefore, the coefficients $-A_{nj}/A_{nn}$ in (A-5) satisfy the (n-1) equations

$$\sum_{j=1}^{n-1} C_{ij}\left(\frac{-A_{nj}}{A_{nn}}\right) = C_{in} \tag{A-7}$$

for $i = 1, 2, \ldots, n-1$. So, when we put

$$x_n = \sum_{j=1}^{n-1} \gamma_j x_j \quad , \tag{A-8}$$

the $(\gamma_1, \ldots \gamma_{n-1})$ vector is the unique solution to the matrix equation equivalent to (A-7):

$$C \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{n-1} \end{bmatrix} = \begin{bmatrix} C_{1n} \\ \vdots \\ C_{n-1,n} \end{bmatrix} \tag{A-9}$$

where C is the (n-1)-rowed matrix with elements C_{ij}.

The equation (A-9) can be solved recursively in a simple manner. Indeed, let $(\gamma'_1, \ldots, \gamma'_{n-1}, \gamma'_n)$ be the corresponding coefficients when n is replaced by n+1. Then we have to solve the equation

$$\left[\begin{array}{c|c} C & \begin{matrix} C_{1n} \\ \vdots \\ C_{n-1,n} \end{matrix} \\ \hline C_{n1} \cdots C_{n,n-1} & C_{nn} \end{array}\right] \left[\begin{array}{c} \gamma'_1 \\ \vdots \\ \gamma'_{n-1} \\ \hline \gamma'_n \end{array}\right] = \left[\begin{array}{c} C_{1,n+1} \\ \vdots \\ C_{n-1,n+1} \\ \hline C_{n,n+1} \end{array}\right] . \tag{A-10}$$

We observe that

$$\begin{bmatrix} C_{1,n+1} \\ \vdots \\ C_{n-1,n+1} \end{bmatrix} = \rho \begin{bmatrix} C_{1,n} \\ \vdots \\ C_{n-1,n} \end{bmatrix} \tag{A-11}$$

using (A-1). So the first n-1 rows of (A-10) can be written

$$C \begin{bmatrix} \gamma_1' \\ \vdots \\ \gamma_{n-1}' \end{bmatrix} = -\gamma_n' \begin{bmatrix} C_{1n} \\ \vdots \\ C_{n-1,n} \end{bmatrix} + \rho \begin{bmatrix} C_{1n} \\ \vdots \\ C_{n-1,n} \end{bmatrix} , \tag{A-12}$$

and when this is compared with (A-9), we see that

$$\begin{bmatrix} \gamma_1' \\ \vdots \\ \gamma_{n-1}' \end{bmatrix} = (\rho - \gamma_n') \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{n-1} \end{bmatrix} , \tag{A-13}$$

and this shows how the γ_j' are obtainable from γ_j for $j \le n-1$, once γ_n' is known. The last row of (A-10) is now used, to give

$$(C_{n1},\ldots,C_{n,n-1}) \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{n-1} \end{bmatrix} (\rho-\gamma_n') + \gamma_n' C_{nn} = C_{n,n+1} \tag{A-14}$$

in which (A-13) has been used. This equation yields

$$\gamma_n' = \frac{C_{n,n+1} - \rho \sum_{j=1}^{n-1} C_{nj}\gamma_j}{C_{nn} - \sum_{j=1}^{n-1} C_{nj}\gamma_j} . \tag{A-15}$$

It is convenient here to introduce the quantities

$$u_{n-1} = \sum_{j=1}^{n-1} C_{nj}\gamma_j . \tag{A-16}$$

Then (A-15) becomes

$$\gamma'_n = \frac{\rho\sigma_A^2 - \rho u_{n-1}}{\sigma_A^2 + \sigma_B^2 - u_{n-1}} \qquad \text{(A-17)}$$

making use of (A-1). Finally, replacing n by n+1 in (A-16) yields

$$u_n = \sum_{j=1}^{n} C_{n+1,j}\gamma'_j$$

$$= \rho \sum_{j=1}^{n-1} C_{n,j}\gamma_j(\rho - \gamma'_n) + \rho\sigma_A^2\gamma'_n$$

where we have used (A-11), (A-13), and (A-1). With (A-16), we now can write

$$u_n = \rho(\rho - \gamma'_n)u_{n-1} + \rho\sigma_A^2\gamma'_n \quad . \qquad \text{(A-18)}$$

The recursive solution for the γ's is now complete. But the γ's themselves are not what we want; it is the $\hat{x}$'s , and the γ's are merely a means to an end.

Suppose then that we have obtained $\hat{x}_n$, and u_{n-1} . From (A-8), with n+1 in place of n , we get

$$\hat{x}_{n+1} = \sum_{j=1}^{n-1} \gamma'_j x_j + \gamma'_n x_n \quad . \qquad \text{(A-19)}$$

We use (A-13) now to write

$$\hat{x}_{n+1} = \sum_{j=1}^{n-1} \gamma_j x_j (\rho - \gamma'_n) + \gamma_n x_n \ ,$$

i.e.,

$$\hat{x}_{n+1} = (\rho - \gamma'_n)\hat{x}_n + \gamma'_n x_n \qquad \text{(A-20)}$$

This equation permits the recursive calculation of $\hat{x}_n$ when combined with (A-17) and (A-18). The starting values are

$$\left.\begin{aligned} x_1 &= 0 \\ u_o &= 0 \end{aligned}\right\} \qquad \text{(A-21)}$$

Finally, the σ_n^2 of (A-3) are obtainable from

$$\sigma_n^2 = E(x_n^2) - E(\hat{x}_n^2) = \sigma_A^2 + \sigma_B^2 - (\gamma_1, \ldots, \gamma_{n-1}) \begin{bmatrix} C_{1n} \\ \vdots \\ C_{n-1,n} \end{bmatrix} ,$$

i.e.,

$$\sigma_n^2 = \sigma_A^2 + \sigma_B^2 - u_{n-1} \quad . \tag{A-22}$$

The only equations required for the actual computations are (A-17), (A-18), (A-20), (A-21), and (A-22).

APPENDIX B

The following 7 tables display the estimates that resulted from 80 independent samples of time series $x_1, \ldots, x_{100}$ (miss distances in a burst of 100 rounds), simulated by the computer. The heading of each table defines the quantity whose values are tabulated in the body of the table. The 4 numbers on the left and the 4 at the top represent, respectively, the true value of ρ (or ζ) and σ_A. The value of σ_B was taken to be 2.0 throughout. For each σ_A, ρ there are 5 numbers presented, which correspond to the 5 independent samples used for each parameter combination. See the main text for additional details.

As an example, the estimates for σ_B, viz., $E(\sigma_B)$, for $\rho = .80$ and $\sigma_A = 4$ were 1.4, 1.8, 1.8, 2.3, 2.3.

The parameters ρ and ζ are related to each other by the equation

$$\rho = \tanh(\zeta).$$

We used ζ in the computer calculations.

TABLE B-1

$E[\sigma_A]$

	σ_A			
ρ	6	4	2	1
.98	4.4	7.0	2.8	2.2
	8.6	7.2	3.4	1.6
	10.7	4.8	1.8	1.0
	3.9	5.0	2.1	1.1
	8.7	6.3	1.7	1.8
.90	10.7	3.8	2.6	1.6
	4.8	5.7	2.1	1.2
	4.4	7.6	1.9	0.9
	7.1	4.7	2.1	1.4
	5.1	5.0	2.0	1.7
.80	5.4	4.9	2.2	1.6
	8.8	3.8	3.4	0.9
	7.9	6.5	1.8	1.1
	4.9	5.2	2.3	0.8
	7.0	7.2	3.4	2.4
.50	5.5	4.0	2.2	2.4
	5.3	3.8	1.8	1.1
	5.0	3.9	2.0	1.6
	4.7	4.1	1.2	0.6
	6.4	4.3	1.9	0.7

TABLE B-II

$E[\sigma_B]$

ρ	σ_A			
	6	4	2	1
.98	2.5	1.7	2.2	1.8
	2.2	2.2	2.2	2.0
	1.5	2.2	2.0	2.1
	1.7	2.1	1.8	1.5
	2.2	2.1	1.9	2.0
.90	2.4	2.6	2.2	2.1
	1.4	2.3	2.2	1.9
	2.2	1.7	1.9	1.9
	2.5	2.0	2.0	2.1
	2.3	2.3	1.7	1.9
.80	1.4	1.4	1.9	2.2
	1.3	1.8	1.6	1.7
	1.7	1.8	2.2	1.8
	2.1	2.3	2.1	2.3
	2.2	2.3	2.4	2.2
.50	2.6	1.2	1.8	1.1
	2.2	2.8	2.3	2.4
	4.3	2.3	1.9	1.2
	3.2	2.3	1.9	1.9
	2.8	1.4	2.0	2.1

TABLE B-III

$\hat{\rho} = \tanh(\hat{\zeta})$ (where $\hat{\zeta} = [\zeta]$)

	^{n}A			
ρ	6	4	2	1
.98	.99	.99	.98	.98
	.98	.99	.99	.98
	.99	.99	.96	.99
	.89	.98	.97	.75
	.99	.99	.98	.98
.90	.97	.97	.93	.97
	.76	.94	.96	.86
	.80	.96	.94	.95
	.96	.92	.88	.97
	.85	.97	.81	.95
.80	.68	.84	.85	.82
	.90	.85	.78	.82
	.86	.91	.71	.72
	.70	.85	.89	.92
	.91	.91	.95	.94
.50	.47	.39	.66	.53
	.48	.65	.76	.93
	.83	.84	.62	.57
	.55	.68	.66	.82
	.62	.62	.80	.94

TABLE B-IV

$\hat{\zeta} = E[\zeta]$

ζ	σ_A			
	6	4	2	1
2.30	2.59	2.78	2.36	2.43
	2.43	2.60	2.48	2.40
	2.50	2.57	2.01	2.48
	1.43	2.30	2.04	0.97
	2.72	2.70	2.25	2.33
1.47	2.11	2.16	1.64	2.09
	1.01	1.72	1.98	1.28
	1.09	2.01	1.78	1.84
	1.98	1.61	1.38	2.04
	1.27	2.07	1.13	1.79
1.10	0.83	1.23	1.26	1.15
	1.45	1.26	1.05	1.16
	1.28	1.51	0.88	0.91
	0.86	1.24	1.40	1.58
	1.56	1.54	1.84	1.78
.55	0.51	0.41	0.79	0.59
	0.52	0.78	0.99	1.66
	1.19	1.23	0.72	0.65
	0.62	0.84	0.79	1.16
	0.72	0.72	1.09	1.72

TABLE B-V

$$[D^2(\sigma_A)]^{\frac{1}{2}} = \left\{E[(\sigma_A - \hat{\sigma}_A)^2]\right\}^{\frac{1}{2}}$$

	σ_A			
ρ	6	4	2	1
.98	1.24	1.42	.50	.95
	2.54	3.47	1.03	.63
	1.30	.70	.72	.54
	.92	1.34	.88	.61
	1.33	1.47	.69	.70
.90	2.09	1.00	1.09	.93
	1.01	1.67	1.10	.66
	.86	1.60	.71	.51
	1.87	1.25	.58	.65
	1.35	1.46	.59	.80
.80	.69	1.38	.59	.67
	1.78	1.09	.85	.60
	1.93	1.31	.75	.62
	.78	1.26	.79	.64
	1.37	2.36	1.08	1.06
.50	.95	.69	.44	.47
	1.10	1.23	.84	.65
	1.50	.99	.90	.79
	1.07	1.02	.71	.48
	1.19	.96	.72	.61

TABLE B-VI

$$\left[D^2(\sigma_B)\right]^{\frac{1}{2}} = \left\{E\left[(\sigma_B - \hat{\sigma}_B)^2\right]\right\}^{\frac{1}{2}}$$

	σ_A			
ρ	6	4	2	1
.98	.22	.15	.20	.16
	.22	.21	.16	.18
	.12	.24	.18	.16
	.34	.16	.20	.48
	.18	.21	.22	.13
.90	.61	.32	.27	.40
	.87	.35	.29	.33
	.55	.35	.41	.24
	.36	.31	.35	.16
	.45	.25	.63	.21
.80	.52	.68	.49	.45
	.68	.50	.59	.62
	1.03	.47	.53	.43
	.84	.82	.26	.28
	.33	.43	.23	.22
.50	1.13	.82	.36	.55
	1.81	1.39	.63	.38
	.78	1.05	.83	.88
	1.26	1.11	.64	.30
	1.59	1.18	.76	.24

TABLE B-VII

$$\left[D^2(\zeta)\right]^{\frac{1}{2}} = \left\{E\left[(\zeta-\hat{\zeta})^2\right]\right\}^{\frac{1}{2}}$$

	σ_A			
ζ	6	4	2	1
2.30	.36	.21	.27	.41
	.28	.42	.25	.38
	.09	.26	.38	.50
	.32	.28	.44	.78
	.10	.39	.54	.46
1.47	.38	.38	.61	.75
	.27	.36	.70	.79
	.32	.39	.78	.81
	.31	.23	.38	.51
	.33	.36	.63	.60
1.10	.18	.32	.57	.67
	.23	.54	.43	.35
	.35	.28	.69	.64
	.25	.36	.41	.84
	.25	.51	.49	.50
.55	.15	.27	.30	.25
	.21	.37	.71	.86
	.54	.68	.58	.71
	.41	.33	.70	.82
	.34	.33	.82	1.02