Waypoint Analysis for Command and Control

David A. Wendt,[1] Mark E. Irwin,[2] Noel Cressie[2]

[1] *Battelle Memorial Institute, Columbus, Ohio 43201*

[2] *Department of Statistics, The Ohio State University, Columbus, Ohio 43210*

Received 12 July 2002; revised 28 May 2004; accepted 9 June 2004
DOI 10.1002/nav.20042
Published online 25 August 2004 in Wiley InterScience (www.interscience.wiley.com).

Abstract: Command and Control (C2) in a military setting can be epitomized in battles-of-old when commanders would seek high ground to gain superior spatial-temporal information; from this vantage point, decisions were made and relayed to units in the field. Although the fundamentals remain, technology has changed the practice of C2; for example, enemy units may be observed remotely, with instruments of varying positional accuracy. A basic problem in C2 is the ability to track an enemy object in the battlespace and to forecast its future position; the (extended) Kalman filter provides a straightforward solution. The problem changes fundamentally if one assumes that the moving object is headed for an (unknown) location, or waypoint. This article is concerned with the new problem of estimation of such a waypoint, for which we use Bayesian statistical prediction. The computational burden is greater than an ad hoc regression-based estimate, which we also develop, but the Bayesian approach has a big advantage in that it yields both a predictor and a measure of its variability. © 2004 Wiley Periodicals, Inc. Naval Research Logistics 51: 1045–1067, 2004.

Keywords: battlespace; Bayesian analysis; C2; spatio-temporal prediction; hierarchical statistical model

1. INTRODUCTION

The topics known as C2 (meant as an abbreviation for "Command and Control") and its intellectual brethren C3, C4I, etc., are umbrella terms used to describe a body of research and applications dedicated to incorporating new technologies into all branches of the armed forces. Applications incorporated into the greater C2 framework include, but are not limited to, command applications, operations applications, intelligence applications, fire-support applications, logistics applications, and communications applications (e.g., Paté-Cornell and Fischbeck [6]; Sherrill and Barr [7]; Wendt, Cressie, and Johanneson [9]). All branches of the armed services share C2 needs. Consequently, applications should be sufficiently flexible to work across different services and different members of allied forces, as needed.

Correspondence to: N. Cressie (ncressie@stat.ohio-state.edu)

Statisticians have a unique perspective on the challenges presented by C2 and thus have an opportunity to contribute to the research and applications in this area. Some of the areas that may be of interest to commanders at all levels are statistical methodology for situational and predictive battlespace awareness, change and anomaly detection, spatial-temporal prediction, measures of information and understanding, and decision theory. Intelligence analysts are interested in developing methods to address confidence, accuracy, and completeness of source data, the validity of information-processing results, and the quantification of uncertainty or confidence in statistical results.

In this article, we consider a very simple situation, namely inference (estimation or prediction) on an object in the battlespace based on noisy, incomplete data. One commonly explored collection of the object's properties is its current location and, perhaps, its movement vector, which leads to target tracking (e.g., see the February 2002 issue of *IEEE Transactions on Signal Processing*). Rather than predicting the current and short-term future state of the object, in this article we shall concentrate on its final destination. We assume that the mobile object is attempting to reach a predesignated location at a predesignated time, namely, a *waypoint,* and it is this quantity that we wish to make inference on. To the best of our knowledge, this problem has not been addressed in the statistical forecasting literature.

Formally, a waypoint is a point (W_x, W_y) in space that a mobile object is attempting to reach at a time W_τ. Military commanders on both sides may use collections of waypoints to distribute orders to units and to coordinate their forces. Often a series of waypoints, called a *phaseline,* is used to describe more complicated instructions. In this article, we consider a single object and define $\mathbf{W} = (W_x, W_y, W_\tau)$ to be its waypoint.

In all that is to follow, we make the following simplifying assumptions (whose generalizations are discussed in Section 7). We recognize that we are making first steps in a problem that could become easily complicated, and hence we make no claim that we have a solution that is ready for use in the field. We assume a flat landscape, so that terrain effects can be ignored. We also assume that paths are defined by a single waypoint. We recognize that multiple waypoints can be used to define a more complicated path, but the basic building block is an object headed for a single destination. It is implicit in our equations of object movement that the object makes no drastic detours. For example, the "object" is a boat, the "terrain" an open body of water, the "movement equations" are due to the boat's being unmanned, and the "waypoint" is a mine-drop location. For situations where the data are updated in discrete time, it is natural to develop the model and methodology likewise; here, we assume discrete time points, $1, 2, \ldots, W_\tau$.

When modeling the movement of an object in space, two of the primary considerations are the choice of coordinate system and the type of process (or state) error associated with the object's movement. Two sets of coordinate systems present themselves as most convenient for modeling location or movement. Standard Cartesian coordinates, denoted (x, y), are easily understood and manipulated. On the other hand, polar coordinates (radius, angle), can be convenient when doing analysis from a fixed frame of reference, and are most natural when modeling an object's speed and angle. We choose to work mainly, although not exclusively, with polar coordinates in this article.

Differentiation of the position function of an object provides a series of choices for process error. Process error on the object's position can be a simple and convenient choice, but it can be hard to disentangle the observational error. A second option is to put the process error on the components of velocity. It seems intuitive to think of speed and change of angle when considering movement toward a waypoint. Further differentiation of speed yields acceleration, which is also a possible place to model process error. In Section 2.2, we choose to place error on the components of velocity (speed and angle), thereby compromising between the simplicity of position error and the more complicated but perhaps more realistic acceleration error.

Based on the discussion above, we define the process model in Section 2 (and Appendix A). Upon combining this model with a model for the observational (or measurement) error and a prior distribution of the waypoint, Bayesian methods applied to the state and measurement equations give a posterior distribution of the waypoint. For the purposes of comparison, we also develop an ad hoc, regression-based estimate of the waypoint. One can imagine a collection of vectors describing an object's future path generated by connecting any two observed positions of the object; then least-squares triangulation leads to simple linear regression and an estimate of the waypoint.

This article is organized as follows. Section 2 reviews Bayesian space-time hierarchical modeling, particularly the development of state and measurement equations. The state equations, or process model, are guided by physical properties of motion, while the measurement equations, or data model, describe error in the measurement system. Section 2 also considers the parameter model, often known as the prior distribution. In Section 3, the Bayesian hierarchical method of inference on the waypoint is considered, along with an ad hoc, regression-based approach. The predictive distributions are not available in closed form; Section 4 gives a Markov chain Monte Carlo (MCMC) method that allows any functional of the state process to be predicted. Section 5 addresses the methodology used to simulate realistic waypoint data, and Section 6 describes the simulation study conducted to compare the Bayesian and regression-based estimators. In Section 7, we discuss our results and consider future directions.

2. SPACE-TIME HIERARCHICAL MODELING

In this section, we describe the basic principles of Bayesian hierarchical modeling and how those principles might be applied to the estimation of waypoints. Consider the state $\{\mathbf{X}_t : t \in T \subset \mathbb{R}\}$ of the object of interest, where $\mathbf{X}_t \equiv (x_t, y_t, s_t, \theta_t)$. Here (x_t, y_t) represents the true position of the object in a standard Cartesian coordinate system at time t, and (s_t, θ_t) represents the speed and angle components of the object's instantaneous velocity at time t, where we recall from Section 1 that $t = 1, 2, \ldots, W_\tau$. Thus, $\mathbf{X}_t$ describes the movement state of an object at time t, of which we have imperfect knowledge through noisy current and past data, $\mathbf{Z}_{1:t} \equiv (\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_t)$. For example, $\mathbf{Z}_k$ might be a noisy observation of (x_k, y_k) obtained from a radar station or even from an observer with binoculars, $k = 1, \ldots, t$.

The Bayesian statistical approach offers a coherent mechanism for using scientific reasoning, learning from data, and prediction, all in a framework that manages the various sources of uncertainty. The modifier "hierarchical" in the section heading refers to the development of complex stochastic models, built conditionally at various levels. The following summary of hierarchical Bayesian modeling, found in Berliner [2], is useful in describing this framework. The joint statistical distribution for all unknowns is developed as the product of three primary components, namely,

- Data Model:
$$[\mathbf{Z}_{1:t}|\mathbf{X}_{1:t}, \mathbf{\Phi}_1], \qquad t = 1, 2, \ldots, W_\tau,$$
where $\mathbf{X}_{1:t} \equiv (\mathbf{X}_1, \ldots, \mathbf{X}_t)$.
- Process Model:
$$[\mathbf{X}_{1:W_\tau}|\mathbf{W}, \mathbf{\Phi}_2].$$

- Parameter Model:

$$[\mathbf{W}, \mathbf{\Phi}],$$

where $\mathbf{\Phi} \equiv (\mathbf{\Phi}_1, \mathbf{\Phi}_2)$.

Note that $[A|B]$ is standard notation for the conditional distribution of A given B.

The essence of this hierarchical approach is that we separate the modeling of the data $\{\mathbf{Z}_k\}$, and the errors implicit in them, from the modeling of $\{\mathbf{X}_k\}$, the underlying state of interest. Bayes' Theorem provides a mechanism for developing the *posterior distribution* of the unknowns conditional on the observations. Specifically, we are interested in $[\mathbf{W}|\mathbf{Z}_{1:t}]$, and possibly in the object's path, $[\mathbf{X}_{1:W_\tau}|\mathbf{Z}_{1:t}]$, where we note that the unknown parameters $\mathbf{\Phi}$ are integrated out in our fully Bayesian approach (e.g., Gelman et al. [4], Chap. 3). Contrast this with an empirical Bayesian approach where $[\mathbf{W}|\mathbf{Z}_{1:t}, \mathbf{\Phi}]$ and $[\mathbf{X}_{1:W_\tau}|\mathbf{Z}_{1:t}, \mathbf{\Phi}]$ are obtained, and a data-based estimate $\hat{\mathbf{\Phi}}$ is substituted for $\mathbf{\Phi}$ (e.g., Carlin and Louis [3]). The latter approach yields variability estimates that are typically too small.

2.1. Measurement Equations—The Data Model

We alluded to observations made by radar-type or other sensors earlier in this article. For now, we simply let $\mathbf{Z}_k$ represent the observation of an object's position at time k. Then a key statistical modeling task is the specification of the data given the process. An assumption of conditional independence yields

$$[\mathbf{Z}_{1:t}|\mathbf{X}_{1:t}, \mathbf{\Phi}_1] = \prod_{k=1}^{t} [\mathbf{Z}_k|\mathbf{X}_k, \mathbf{\Phi}_1].$$

(Notice that we are not saying that for any two times the $\mathbf{Z}$'s are marginally independent. Indeed, these data are likely to be highly dependent, due to the dependence of the underlying $\mathbf{X}$'s.) The conditional-independence assumption usually holds and is based on the idea that multiple measurements of the same object differ according to errors that are independent. Here the parameters in $\mathbf{\Phi}_1$ represent measurement-error variances and covariances that are known from instrument manufacturers' quality assurance.

2.2. State Equations—The Process Model

The key in this step is to build a stochastic process for

$$[\mathbf{X}_{1:W_\tau}|\mathbf{W}, \mathbf{\Phi}_2].$$

Our basic assumption here is that the object's location at time $(t + 1)$ given $\mathbf{X}_{1:t}$ (and the waypoint $\mathbf{W}$) depends only on the previous state $\mathbf{X}_t$. That is, we assume that

$$[\mathbf{X}_{t+1}|\mathbf{X}_{1:t}, \mathbf{W}, \mathbf{\Phi}_2] = [\mathbf{X}_{t+1}|\mathbf{X}_t, \mathbf{W}, \mathbf{\Phi}_2].$$

As a consequence,

$$[\mathbf{X}_{1:t}|\mathbf{W}, \mathbf{\Phi}_2] = [\mathbf{X}_1|\mathbf{W}, \mathbf{\Phi}_2] \prod_{k=2}^{t} [\mathbf{X}_k|\mathbf{X}_{k-1}, \mathbf{W}, \mathbf{\Phi}_2].$$

This model is new in that the object's destination $\mathbf{W}$ enters explicitly into the manner in which the state $\mathbf{X}_{t+1}$ evolves from $\mathbf{X}_t$. We shall now consider an example of this type of process model.

At each time-point of interest, the object evaluates its position with respect to the waypoint. Based on this information, the speed and angle components of the velocity needed to reach the next waypoint at the desired time can be computed. However, these required values of speed and angle are contaminated with process error, resulting in realized values of speed and angle. This results in a spatial displacement which, when added to the previous location, gives the updated location of the object. Such a model would be appropriate for an unmanned vehicle on a flat terrain (e.g., an unmanned boat navigating an open body of water). More complicated models are possible that incorporate, for example, the possibility of failures en route, rest stops, and other factors such as terrain.

The recursive state equation, based on the calculations in Appendix A, is

$$\mathbf{X}_{t+1} \equiv \begin{pmatrix} x_{t+1} \\ y_{t+1} \\ s_{t+1} \\ \theta_{t+1} \end{pmatrix}$$

$$= \begin{pmatrix} x_t + \dfrac{1}{\theta_{t+1} - \theta_t}\left(s_{t+1}\sin\theta_{t+1} - s_t\sin\theta_t + \dfrac{s_{t+1} - s_t}{\theta_{t+1} - \theta_t}(\cos\theta_{t+1} - \cos\theta_t)\right) \\ y_t + \dfrac{1}{\theta_{t+1} - \theta_t}\left(-s_{t+1}\cos\theta_{t+1} + s_t\cos\theta_t + \dfrac{s_{t+1} - s_t}{\theta_{t+1} - \theta_t}(\sin\theta_{t+1} - \sin\theta_t)\right) \\ \dfrac{\sqrt{(W_x - x_t)^2 - (W_y - y_t)^2}}{W_\tau - t} + \varepsilon_{s,t} \\ \tan^{-1}\left(\dfrac{W_y - y_t}{W_x - x_t}\right) + \varepsilon_{\theta,t} \end{pmatrix}. \quad (1)$$

Notice the explicit dependence of the state equation on the final destination (waypoint) $\mathbf{W}$; clearly, the Kalman filter and its variants cannot be used for inference on the waypoint.

It is worth noting that given the waypoint $\mathbf{W}$, $\mathbf{X}_{t+1}$ can be written as function of $\mathbf{X}_t$ and a random component. In other words, the distribution of $\mathbf{X}_{t+1}$ given $\mathbf{X}_t$ is completely defined by the distribution of the error vector $(\varepsilon_{s,t}, \varepsilon_{\theta,t})$, associated with the speed and angle at time t. Here the parameters $\mathbf{\Phi}_2$ are variances of $\varepsilon_{s,t}$ and $\varepsilon_{\theta,t}$. Knowing how much the next time-point's speed- and angle-values can deviate from those of the current time point, allows $\mathbf{\Phi}_2$ to be specified.

2.3. Prior Specification—The Parameter Model

The final piece of the Bayesian hierarchical model is the prior, or parameter, model. We shall consider $\mathbf{\Phi}$ and $\mathbf{W}$ to be independent,

$$[\mathbf{W}, \mathbf{\Phi}] = [\mathbf{W}][\mathbf{\Phi}],$$

and specify the prior distribution of [**W**] and [**Φ**] separately. Note that this assumption represents a first approximation and one can imagine situations where knowledge of $\mathbf{\Phi}_2$ can be informative about **W**. In that case, we would use the general relation, $[\mathbf{W}, \mathbf{\Phi}] = [\mathbf{W}|\mathbf{\Phi}][\mathbf{\Phi}]$.

The selection, or modeling, of prior distributions depends upon expert knowledge of the parameters, statistical convenience, or modeler intuition. In situations where little is known about the parameters of interest, it is often desirable to choose a prior distribution that is as "flat" as possible (the extreme case being a noninformative prior) (see, for example, Gelman et al. [4], Section 2.8). However, in military settings, there can be information from intelligence sources that results in an informative prior whose probability density function is narrower. See Section 6 for specification of [**W**] in the examples discussed there.

3. WAYPOINT ANALYSIS

Bayesian inference on (W_x, W_y) and W_τ is powerful, since it automatically gives measures of variability through the posterior distribution. However, its implementation may not be straightforward. After presenting the fully Bayesian approach in Section 3.1, an ad hoc regression-based estimate (derived from least-squares triangulation) of **W** is given in Section 3.2. This estimate is useful for comparison to the Bayes estimate, and can be used informally to calibrate the prior on **W**.

3.1. Hierarchical Probability Structure

We are interested in the probability density of the waypoint given the data. Using Bayes' Theorem, we can write

$$[\mathbf{W}, \mathbf{X}_{1:W_\tau}|\mathbf{Z}_{1:t}] \propto [\mathbf{Z}_{1:t}|\mathbf{X}_{1:t}][\mathbf{X}_{1:W_\tau}|\mathbf{W}][\mathbf{W}]. \tag{2}$$

Note that, for ease of notation, the parameters **Φ** do not appear explicitly in our derivations. However, they are implicitly present, and we are cognizant of their importance in the model.

In practice, it may be difficult to put an informative prior [**W**], on the waypoint. In the Bayesian approach, the better the available intelligence, the more informative the prior can be, and hence the better the Bayesian approach will perform. In what follows, we shall specify W_τ and consider a reasonably uninformative (i.e., "flat") prior for (W_x, W_y) given W_τ in the region of interest. Recall the example of the (unmanned) enemy boat loaded with mines and headed for an unknown mine-drop location. Friendly intelligence is aware that the enemy's laying of mines is time-sensitive to the passage of a friendly aircraft carrier, so that it can provide a value of W_τ within a more or less narrow window. Then the problem is to make inference on the mine-drop location (W_x, W_y). When terrain features are incorporated (e.g., unnavigable waters, islands), along with the military commander's intuition, there may be only a few possible waypoint locations, each of which would receive a prior probability based on intelligence analysis.

Formally, we decompose the prior on the waypoint into spatial and temporal components, as follows:

$$[\mathbf{W}] = [W_x, W_y|W_\tau][W_\tau],$$

and then make $[W_\tau]$ a degenerate probability distribution. There is a second way to decompose the prior on the waypoint, namely,

$$[\mathbf{W}] = [W_\tau | W_x, W_y][W_x, W_y],$$

which may be sensible in some battlespace scenarios. This second approach will not be pursued further here.

Given W_τ, the waypoint-location component (W_x, W_y) might be represented using a mixture of bivariate unimodal densities centered at one or more key locations in the battlespace. Although W_τ is specified in advance, we could then vary W_τ to determine the sensitivity of our results to its specification. This is different from putting a prior on W_τ (see the discussion in Section 7).

The next step in the Bayesian approach is to obtain the posterior distribution of all unknowns given the data (and the specified W_τ). In (2), $[\mathbf{Z}_{1:t} | \mathbf{X}_{1:t}]$ is straightforward to calculate from the measurement equation, and $[\mathbf{W}]$ (actually $[W_x, W_y | W_\tau]$ will be specified). Therefore, we focus on $[\mathbf{X}_{1:W_\tau} | \mathbf{W}]$, the conditional probability of the actual states, given the waypoint. Then, from Section 2.2,

$$[\mathbf{X}_{1:t} | \mathbf{W}] = [\mathbf{X}_t | \mathbf{X}_{1:t-1}, \mathbf{W}][\mathbf{X}_{1:t-1} | \mathbf{W}]$$

$$= [\mathbf{X}_1 | \mathbf{W}] \prod_{k=2}^{t} [\mathbf{X}_k | \mathbf{X}_{k-1}, \mathbf{W}],$$

and the latter distributions are available to us from the state equation (1).

In the Bayesian analysis that is to follow, we shall condition the whole analysis on W_τ and define $\mathbf{W}_\mathbf{s} \equiv (W_x, W_y)$ to be the unknown spatial location of the waypoint. In order to generate the posterior distribution of the waypoint given the observed path, we propose using a Markov chain Monte Carlo (MCMC) algorithm called the Gibbs sampler (see, e.g., Gelman et al. [4], Chap. 11). This will require sampling iteratively from the path given the data and the waypoint,

$$[\mathbf{X}_{1:W_\tau} | W_\tau, \mathbf{Z}_{1:t}, \mathbf{W}_\mathbf{s}], \tag{3}$$

from the waypoint given the data and the path,

$$[\mathbf{W}_\mathbf{s} | W_\tau, \mathbf{Z}_{1:t}, \mathbf{X}_{1:W_\tau}], \tag{4}$$

and then repeating the sequence until the resulting sampling distribution converges. More details are given in Section 4.

The result of the Gibbs sampler is a set of realizations, $\{\mathbf{W}_\mathbf{s}^{(\ell)}, \mathbf{X}_{1:W_\tau}^{(\ell)} : \ell = 1, \ldots, L\}$, sampled from the posterior distribution of $\mathbf{W}_\mathbf{s}, \mathbf{X}_{1:W_\tau}$, given W_τ and $\mathbf{Z}_{1:t}$. This automatically yields L realizations of any functional. For example, assuming squared-error loss, the best estimate of the waypoint location $\mathbf{W}_\mathbf{s}$, based on current and past data, is $E(\mathbf{W}_\mathbf{s} | W_\tau, \mathbf{Z}_{1:t})$, which is estimated by

$$\mathbf{W}_\mathbf{s}^* \equiv \frac{1}{L} \sum_{\ell=1}^{L} \mathbf{W}_\mathbf{s}^{(\ell)}. \tag{5}$$

Likewise, optimal smoothing, filtering, and forecasting of the path $\mathbf{X}_{1:W_\tau} = (\mathbf{X}_{1:t-1}, \mathbf{X}_t, \mathbf{X}_{t+1:W_\tau})$ is obtained from,

$$\mathbf{X}^*_{1:W_\tau} \equiv \frac{1}{L} \sum_{\ell=1}^{L} \mathbf{X}^{(\ell)}_{1:W_\tau}. \tag{6}$$

Associated measures of variability, like pointwise posterior standard deviations of the elements of $\mathbf{W_s}$, $\mathbf{X}_{1:W_\tau}$, are obtained from analogous formulas based on the L realizations from the Gibbs sampler.

3.2. Regression-Based Estimate Derived from Least-Squares Triangulation

For the purposes of comparison to the Bayes estimate, we now develop an ad hoc estimate of the waypoint that is derived from the principles of triangulation; that is, we find a position by means of bearings from fixed points that are a known distance apart. While the imperfect observation of an object moving along a path does not allow one to obtain the exact position of the object, a statistical triangulation procedure can be considered.

If an object moves along a straight path without error toward the waypoint, then the vector formed by connecting two points on the path will ultimately intersect the waypoint. When error is included in the movement (state) process, the vector connecting any two points can be extended to yield a value equal to the waypoint plus an error term. This leads to a regression-based estimate.

More specifically, consider any two spatial-temporal observations (x_i, y_i, i) and (x_j, y_j, j); $i < j$, of an object moving along a path. From the two spatial locations (x_i, y_i), (x_j, y_j), the line that connects them is the locus of points,

$$\ell_{ij}(v) = \begin{pmatrix} x_j + (v-j)\dfrac{x_j - x_i}{j-i} \\ y_j + (v-j)\dfrac{y_j - y_i}{j-i} \\ v \end{pmatrix}, \qquad v \geq 0.$$

Then write $\ell_{ij}(W_\tau) = (W_x, W_y, W_\tau) + (\varepsilon_x, \varepsilon_y, 0)$. Thus, observing a pair of observations from the path is equivalent to observing a vector that passes through the waypoint plus an error term.

Given the set of such observed lines, $\{\ell_{ij}(\cdot) : i < j\}$, we can consider a least-squares solution to identify (W_x, W_y). That is, for a given W_τ, solve the least-squares equation,

$$(\hat{W}_x, \hat{W}_y) = \arg\min_{\mathbf{W_s}} \sum_{i<j} (\ell_{ij}(W_\tau) - \mathbf{W})'(\ell_{ij}(W_\tau) - \mathbf{W}). \tag{7}$$

After completing the squares and differentiating with respect to W_x and W_y, the estimating equations are:

$$\sum_{i<j} \left(x_j + (W_\tau - j)\frac{x_j - x_i}{j-i} - W_x \right) = 0,$$

$$\sum_{i<j}\left(y_j + (W_\tau - j)\frac{y_j - y_i}{j - i} - W_y\right) = 0.$$

It is a simple matter to solve these equations, yielding

$$\hat{W}_x = \sum_{i<j}\left(x_j + (W_\tau - j)\frac{x_j - x_i}{j - i}\right)\Big/\sum_{i<j} 1 \tag{8}$$

$$\hat{W}_y = \sum_{i<j}\left(y_j + (W_\tau - j)\frac{y_j - y_i}{j - i}\right)\Big/\sum_{i<j} 1. \tag{9}$$

In practice, the locations (x_i, y_i) and (x_j, y_j) are not known precisely and have to be estimated from the data $\mathbf{Z}_{1:t}$. Therefore, replace them with

$$(\hat{x}_k(\mathbf{Z}_{1:t}), \quad \hat{y}_k(\mathbf{Z}_{1:t})), \qquad k = 1, \ldots, t, \tag{10}$$

where $\hat{x}_k$ and $\hat{y}_k$ are ad hoc estimates of the object's coordinates at time k. For example, if there is only one radar station tracking the object, yielding data (Z_{k1}, Z_{k2}) at time k, an obvious estimate would be $\hat{x}_k = Z_{k1}$, $\hat{y}_k = Z_{k2}$. A (weighted) average could be used when there are observations from more than one radar station (weights would be inversely proportional to measurement-error variances).

The regression-based estimate represented by (8) and (9) will be compared to the Bayes estimate (5) by simulation in Section 6. However, the regression-based estimate has no valid measure of variability and in that sense it is uncompetitive with the Bayes estimate.

4. MCMC SCHEME FOR WAYPOINT ANALYSIS

As was seen in Section 3.1, realizations from the posterior distribution $[\mathbf{W_s}, \mathbf{X}_{1:W_\tau} | W_\tau, \mathbf{Z}_{1:t}]$ can be generated by the following Gibbs sampler (e.g., Gelman et al. [4], Section 11.3):

- Sample the complete path from $[\mathbf{X}_{1:W_\tau} | W_\tau, \mathbf{Z}_{1:t}, \mathbf{W_s}]$.
- Sample the waypoint from $[\mathbf{W_s} | W_\tau, \mathbf{Z}_{1:t}, \mathbf{X}_{1:W_\tau}] = [\mathbf{W_s} | W_\tau, \mathbf{X}_{1:W_\tau}]$.

This sequence is then repeated L times. However, the individual steps cannot all be implemented by sampling directly from the conditional distributions, due to nonconjugacies in the distributions. Instead, a combination of Gibbs and Metropolis-Hastings sampling will be used.

4.1. Sampling the Complete Path

Sampling the complete path will be done with two Gibbs steps based on the range of the observed data.

- Step 1: Sample $\mathbf{X}_{1:t}$ from $[\mathbf{X}_{1:t} | W_\tau, \mathbf{X}_{t+1:W_\tau}, \mathbf{Z}_{1:t}, \mathbf{W_s}] = [\mathbf{X}_{1:t} | W_\tau, \mathbf{X}_{t+1}, \mathbf{Z}_{1:t}, \mathbf{W_s}]$.

  The first step is to sample the portion of the path that is in the range of the data. This will be done one time-point at a time by Gibbs sampling, as follows. Sample $\mathbf{X}_k$ from

$$[\mathbf{X}_k|W_\tau, \mathbf{X}_{k-1}, \mathbf{X}_{k+1}, \mathbf{Z}_k, \mathbf{W_s}] \propto [\mathbf{X}_k|W_\tau, \mathbf{X}_{k-1}, \mathbf{W_s}][\mathbf{X}_{k+1}|W_\tau, \mathbf{X}_k, \mathbf{W_s}][\mathbf{Z}_k|\mathbf{X}_k], \quad 2 \le k \le t,$$

and

$$[\mathbf{X}_1|W_\tau, \mathbf{X}_2, \mathbf{Z}_1, \mathbf{W_s}] \propto [\mathbf{X}_2|W_\tau, \mathbf{X}_1, \mathbf{W_s}][\mathbf{Z}_1|\mathbf{X}_1],$$

for $k = 1$. This is difficult to do directly due to the nonlinearities in the state transitions. However, they can be sampled using the Metropolis-Hastings algorithm (e.g., Gelman et al. [4], Section 11.2) with Gaussian proposal distributions; we choose those Gaussian distributions with means and variances obtained from the sampling distributions given above and using the Scaled Unscented Transformation (van der Merwe et al. [8], Irwin, Cressie, and Johanneson [5]). For a brief description of the Metropolis-Hastings algorithm, see Appendix B.

- Step 2: Sample $\mathbf{X}_{t+1\,:\,W_\tau}$ from $[\mathbf{X}_{t+1\,:\,W_\tau}|W_\tau, \mathbf{X}_{1\,:\,t}, \mathbf{W_s}] = [\mathbf{X}_{t+1\,:\,W_\tau}|W_\tau, \mathbf{X}_t, \mathbf{W_s}]$.

  The second step is to sample the future part of the path, which can be done exactly due to the Markov nature of the state process. Simply start the path at the $\mathbf{X}_t$ given in Step 1, and use (1) to generate its evolution to $\mathbf{W_s}$ during the time interval of length $W_\tau - t$. Note that if interest is solely in the posterior distribution of the waypoint, this step is not needed. Sampling from the posterior distribution of the waypoint is detailed in the next subsection.

4.2. Sampling the Waypoint

The form of the desired conditional waypoint distribution is

$$[\mathbf{W_s}|W_\tau, \mathbf{X}_{1\,:\,t}] \propto [\mathbf{W_s}|W_\tau][\mathbf{X}_1|W_\tau, \mathbf{W_s}] \prod_{k=2}^{t} [\mathbf{X}_k|W_\tau, \mathbf{X}_{k-1}, \mathbf{W_s}].$$

Similar to the sampling of the initial part of the path, conditional on the data (Step 1 in Section 4.1), this distribution does not have a tractable form for direct sampling. As in that case, a Metropolis-Hastings (M-H) step will be taken; a brief description of the M-H algorithm is given in Appendix B. However, instead of using the Scaled Unscented Transformation as in Section 4.1, the proposal distribution for $\mathbf{W_s}$ in the M-H step will be based on a Gaussian random walk. The proposed transition for $\mathbf{W_s}$, for Gibbs iterate ℓ, is to sample from

$$\mathbf{W}_\mathbf{s}^{(\ell)} \sim N(\mathbf{W}_\mathbf{s}^{(\ell-1)}, \Sigma_p),$$

where $\mathbf{W}_\mathbf{s}^{(\ell-1)}$ is the realization of the waypoint from the previous iteration of the sampler. The proposal distribution's variance matrix Σ_p will be of the form $\sigma_p^2 I$, where σ_p^2 needs to be chosen to allow for good mixing of the Markov chain. If σ_p^2 is too large, the rejection probability in the M-H step will be too large, whereas setting σ_p^2 too small will lead to slow convergence of the chain to the posterior distribution of $\mathbf{W}_s$ (e.g., Gelman et al. [4], pp. 334, 335). Specific choices are discussed in Section 6.2.

5. BATTLESPACE DATA

In practice, the true path of the object is observed with error and possibly incompletely. The error results from noise in the instrument that measures the object's location, and the magnitude of the error depends on the properties of the measurement device. Incomplete observations may occur due to instrument error or occlusions in the line-of-sight.

To evaluate the waypoint analyses proposed in Section 3, we generate artificial data for use in a simulation experiment (Section 6). There are three steps: First, we generate a "true path"; second, we add measurement noise to the path; and third, we sample observations from the noisy path.

5.1. The True Path

The algorithm we used for generation of the true path is defined recursively using (1). Consequently, the generation of such a path is defined by the starting state $\mathbf{X}_1$, the waypoint (W_x, W_y, W_τ), equations of motion that involve random elements, and the variances of those random elements. We have created code in Matlab to generate as many true paths as needed for the simulation experiment.

5.2. Generation of Noise

To generate the noisy data, we make certain assumptions. First, we assume that only the location, but not the speed and angle, of a mobile object can be observed at any given time. Second, we assume that any noise is additive, zero-mean, and distributed as a multivariate Gaussian random variate. Thus, the noisy data observations can be written:

$$\mathbf{Z}_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{X}_t + \boldsymbol{\varepsilon}_t, \qquad t = 1, 2, \ldots,$$

where $\mathbf{Z}_t$ is a two-dimensional vector made up of the observed location at time t, and $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \Sigma)$.

The form of the (2×2) noise-variance matrix Σ can vary quite a lot, depending on the measuring instrument used. The covariance matrix might be assumed simply to be $\sigma^2 I$, where I is the (2×2) identity matrix. In Wendt, Cressie, and Johanneson [9], a more complex Σ was derived for the error of an observation from a fixed radar-like device. The important feature there was the presence of correlation between observed coordinates, which we capture, using a covariance structure of the form,

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix},$$

where $-1 < \rho < 1$.

5.3. Sampling the Noisy Path

In this article, we focus on the problem of having data on the object's location only up to and including time t. It is possible that some censoring of the data, $\mathbf{Z}_{1:t} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_t)$, also occurs. The methods proposed in this article treat the first observed location as the starting point and assume thereafter that there are no gaps in the data up to and including time t.

Focusing on the observations chosen, rather than on those excluded at all times greater than t, reframes this procedure as sampling. We are, in essence, sampling observations of an object's position up to and including time t in order to forecast the waypoint in the future. In our simulations, as in the analysis, the data are available at equally spaced time points.

6. THE SIMULATION STUDY

Simulation of noisy and incomplete data allows comparison of estimated waypoints with the true waypoint (which is, of course, known for simulated data). In order to examine which circumstances allow for better inference on the waypoint, a simulation study, designed as a factorial experiment was performed. The results of the study were analyzed graphically and by a standard Analysis of Variance (ANOVA).

6.1. Experimental Factors

Six factors were investigated in this study.

True Path (Path): A number of true paths were generated, each with a waypoint location of $\mathbf{W}_s = (100, 100)$ and a waypoint time of $W_\tau = 100$. The starting location for each was the origin (0, 0), with a starting speed of $s_0 = 0$ and a starting direction of $\theta_0 = 0$. The random-process error terms in Eq. (1) were both independent Gaussian random variables with mean 0 and variance 0.1. The choice of mean 0 implies unbiased deviations, and after trying several variances, the variance 0.1 gave relative deviations of the type shown in Figure 1. True paths were generated using (1), and three were chosen; path 1 represents a typical deviation (Fig. 1), path 2 an extreme deviation, and path 3 a rather small deviation from the straight line between (0, 0) and (100, 100). These paths were selected to give three levels for this factor that represented the diversity of possible paths.

Measurement-Error Variance (Var): The distribution of the measurement error (noise) is determined by the size of the overall variation and the level of correlation between the x- and y-components of the noise. Two levels for variance were considered: 0.75 (high) and 0.375 (low). These two levels were chosen with consideration for the amount of 'natural' variation in the state process given by expression (1), specifically a signal-to-noise ratio of 1 and 2, respectively.

Measurement-Error Correlation (Corr): To describe dependence between the x- and y-components of noise, correlation coefficients of $\rho = 0$, $\rho = 1/2$, and $\rho = \sqrt{3}/2$ were considered, representing a range of no, low, and medium dependence. Thus, there are three levels for this factor.

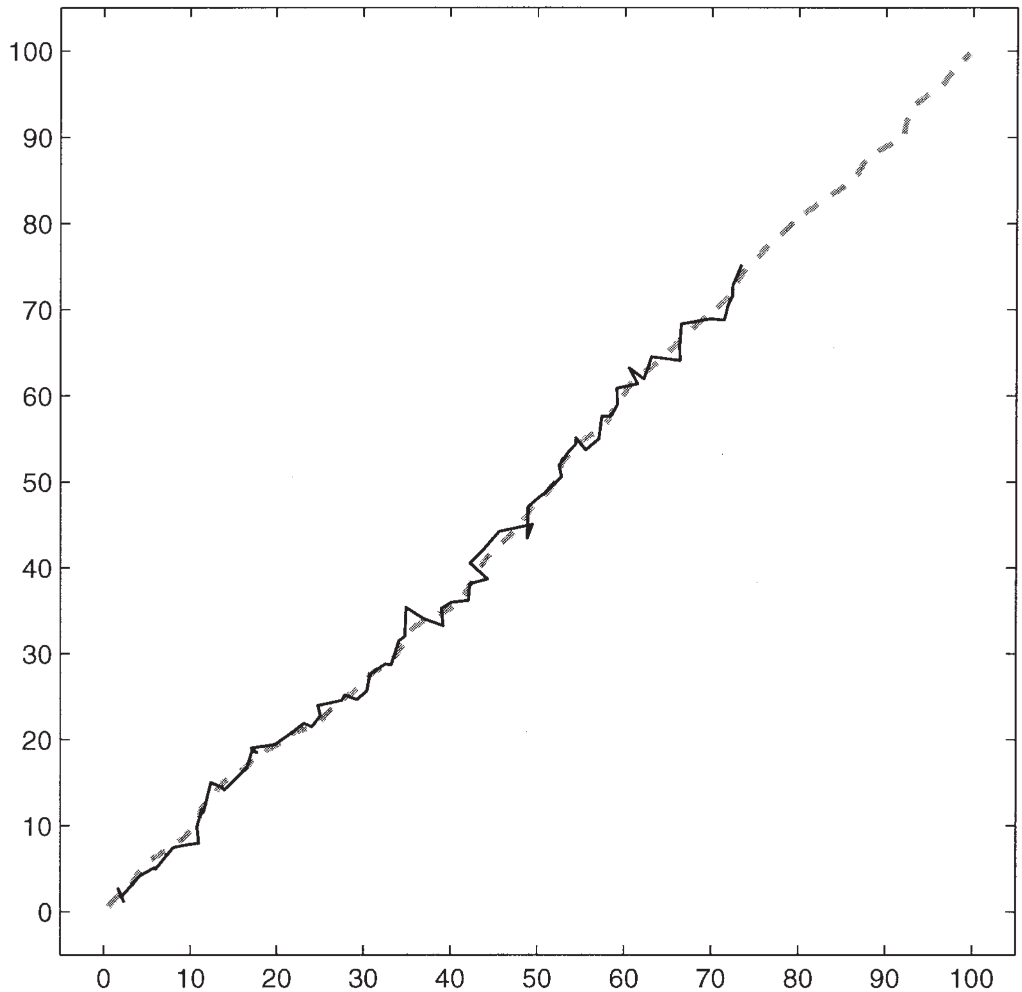

Figure 1. Example of a true path with an observed noisy path. The dashed line is the true path 1. The solid line shows the first 75 observations of path 1, for the high-variance and zero-correlation case.

Sampling of Data (Sample): In this study, 50%, 66%, and 75% of the total possible number of observations were obtained by deleting, respectively, the final 50%, 34%, and 25% of the complete data set. Thus, there are three levels for this factor.

Waypoint Time (WT): By fixing several waypoint times, a sensitivity analysis of the relationship between the spatial and temporal components of the waypoint may be performed. Although the data were generated using a true waypoint time of 100, this is not known in practice, and we must specify a waypoint time as part of the analysis. From a C2 standpoint, this is equivalent to a battle commander forecasting a waypoint location at certain key future times. There are four waypoint times specified, 90, 100, 110, and 120, representing from −10% to 20% of the true waypoint time. Thus, there are four levels for this factor.

Estimation Method: The waypoint is estimated in two different ways in this study. The Bayesian approach is based on the nonlinear state equations given in (1). Using the computationally intensive MCMC methods described in Section 4, the estimate given by (5) is obtained. For comparison, a regression-based estimate is also considered. The resulting estimate, given by (8) and (9), is based on the heuristic that the object always makes a "bee line" for the waypoint. While the regression-based estimate represents a first attempt, it may not do well in situations where the object's movement to the waypoint is more complex. Thus, we can think of the simulation study as comparing two treatments (i.e., the estimates) and we wish to know which treatment is better.

This combination of factors will lead to 18 noisy paths (3 true paths $\times$ 2 variances $\times$ 3 correlations), 54 datasets (3 sampling levels $\times$ 18 noisy paths), and 216 analysis conditions (4 waypoint times $\times$ 54 datasets). Each of these 216 conditions has two waypoint estimates, Bayesian and regression-based. As an example of one of the 54 datasets, Figure 1 shows true path 1 along with the first 75% of the corresponding noisy observed path for the high variance, zero correlation case.

6.2. Other Experimental Parameters

Prior Distribution of the Waypoint: The prior distribution of the waypoint location is Gaussian, with a mean of (100, 100), the true waypoint location, and a variance matrix of $100I$, where I is the (2×2) identity matrix. This distribution describes knowledge about the prior location and prior uncertainty in that location through its mean and variance, respectively. This corresponds to W_x and W_y being independent each with a standard deviation of 10 (i.e., quite an uninformative prior).

MCMC Setup: Each of the MCMC runs has a "burn-in" of 1000 iterations with the next $L = 2000$ iterations used to estimate the waypoint through Eq. (5). This choice was based on standard plots of realizations from the Gibbs sampler versus Markov-chain iteration number (e.g., Carlin and Louis [31]); the chain mixed well after 1000 iterations and the dependence in the chain meant that $L = 2000$ subsequent iterations were needed to obtain a precise estimate given by (5). The variance for the random-walk proposal distribution for sampling $\mathbf{W_s}$ (Section 4.2) was set to $\sigma_p^2 = 5$. This was chosen based on some informal calibration runs that suggested that this choice of σ_p^2 would give a 50% acceptance rate for the Metropolis-Hastings sampler for the waypoint location $\mathbf{W_s}$ (see Gelman et al. [4], p. 335).

The starting state for the sampler was based on the desired path for each object. The location at time t was set to $(x_t, y_t) = (t, t)$, based on a straight-line path between start location and waypoint-prior-mean location; for the same reason, the speed was set to $s_t = \sqrt{2}$ and the direction was set to $\theta_t = \pi/4$; $t = 1, \ldots, 100$. The MCMC starting value for the waypoint was set to the waypoint prior mean (100, 100).

6.3. Computer Implementation

All computer code, for true and noisy path generation, the MCMC sampler for the Bayesian estimation, and for the regression-based estimation, were implemented in Matlab. All analyses were performed on Red Hat Linux 7.3 servers with dual AMD Athlon 1800+ MP CPUs running at 1.533 GHz with 3 GB PC2100 ECC DDR RAM.

It is important to consider the impact of processing time on these analyses. For the settings used in the simulation study described in Section 6, the computation time for the MCMC analysis ranged from 16 minutes when using 50% of the observations to 24 minutes when using 75% of the observations. There is infinitesimal variability in these computation times, due to the fixed truncation of the MCMC. We note that the regression-based estimate is obtained almost instantaneously, but we keep in mind that no valid measure of variation is available. The tradeoff between computing time and accuracy is discussed further in Section 6.6.

6.4. Response Variables

Three response variables were used to examine the results from the simulation study. The primary response of interest is the distance from the estimated waypoint to the true waypoint,

$$d(\hat{\mathbf{W}}_{\mathbf{s}}, \mathbf{W}_{\mathbf{s}}) = \sqrt{(\hat{W}_x - W_x)^2 + (\hat{W}_y - W_y)^2}.$$

The analysis of this response variable will focus on whether the Bayesian estimation procedure gives more accurate estimates of the waypoint than the regression-based estimation procedure. The effects of the other five factors, and their interactions, are also of interest.

The estimated waypoint $\hat{\mathbf{W}}_{\mathbf{s}}$ is the second response of interest. The analysis of this response focuses on the effects of making a correct choice of waypoint time. The relationship between $\hat{\mathbf{W}}_{\mathbf{s}}$ and the five potential descriptive factors (Path, Var, Corr, Sample, and WT) will be examined separately for the Bayesian and the regression-based estimates.

The final response of interest is based on $\Sigma_W \equiv var(\mathbf{W}_{\mathbf{s}}|\mathbf{Z}_{1:t})$, the posterior variance of the waypoint from the Bayesian analysis. The actual response variable is $\log|\Sigma_W|$, where $|\Sigma_W|$ denotes the determinant of Σ_W. The motivation for selecting this response variable is based on the fact that, for a Gaussian distribution, the area inside an equal-density contour is proportional to $\sqrt{|\Sigma_W|}$. Since the posterior distribution of $\mathbf{W}_{\mathbf{s}}$ is approximately Gaussian, this is a useful measure of the amount of information about $\mathbf{W}_{\mathbf{s}}$ in the observed data. The log transform (of the determinant) was used to stabilize the variance and symmetrize the distribution of the residuals from the linear model describing the relationship between $|\Sigma_W|$ and the five potential descriptive factors.

6.5. Results

We first describe the results based on the response $d(\hat{\mathbf{W}}_{\mathbf{s}}, \mathbf{W}_{\mathbf{s}})$. Of the six factors considered in the simulation study, all but Corr have an effect on the size of the estimation error. Estimation Method is extremely important; for 195 of the 216 conditions studied (90%), the estimation error d was smaller for the Bayesian estimate (see also Fig. 5). Section 6.6 discusses this in greater depth, and Figure 2 is illustrative of the superior accuracy of the Bayesian estimate for a particular combination of factor levels. In addition, Estimation Method interacts with the other four factors. Consequently, the outcomes for these four factors need to be examined separately for the Bayesian and for the regression-based estimates. Tables 1 and 2 are formal ANOVA tables that summarize the importance of different factors. (It is not appropriate to declare F-values in the table significant, since our simulation experiment has no randomness in the response variables.)

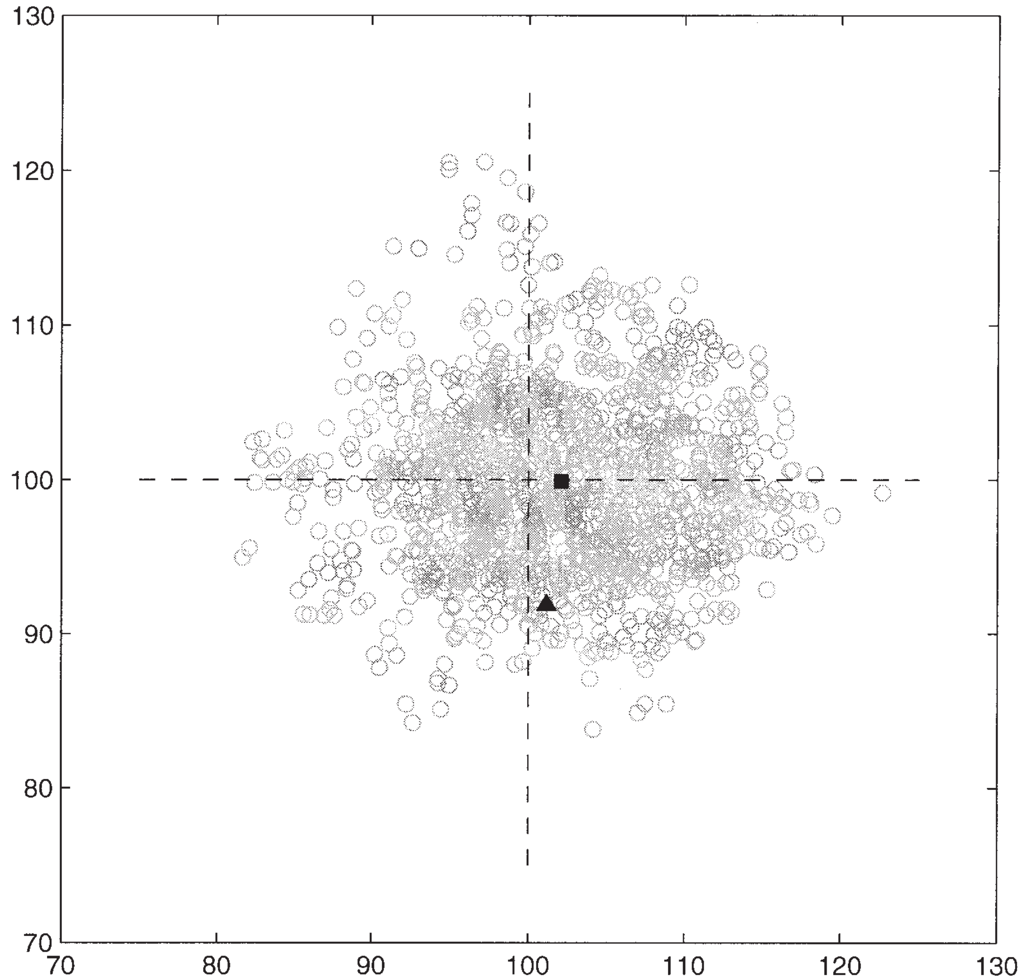

Figure 2. Comparison of Bayesian and regression-based estimates for path 1, low variance, zero correlation, 50 observation, and specified waypoint time = 100. The gray circles are the 2000 Monte Carlo waypoint–location realizations, the square is the Bayesian waypoint–location estimate, and the triangle is the regression-based waypoint–location estimate. The dashed lines intersect at the true waypoint location (100, 100).

For the Bayesian estimation procedure (Table 1), three two-way interactions (Path $*$ WT, Path $*$ Sample, and WT $*$ Sample) are important. The most interesting of these is the WT $*$ Sample interaction. When the correct waypoint time is specified, increasing the percentage of data observed tends to decrease the estimation error. However, when an incorrect waypoint time is specified, the estimation error increases with an increasing percentage of data. This indicates that a bad start cannot be improved by more data. Var is also an important factor, with the smaller estimation errors tending to occur with larger measurement-error variances. This is explained by the fact that the prior will tend to have a stronger effect in the high measurement-error-variance cases, pulling the estimate towards the prior mean.

The results for the regression-based estimation procedure (Table 2) are similar, except that Var does not appear to be important in this case. In addition, while the same two-way

Table 1. Analysis of Variance (ANOVA) table of estimation errors for the Bayesian waypoint-location estimate.

Factor	*Df*	Sum of Sq	Mean Sq	*F* Value
WT	3	5368.833	1789.611	749.062
Path	2	754.266	377.133	157.853
Sample	2	148.623	74.311	31.104
Var	1	63.641	63.641	26.638
Corr	2	4.528	2.264	0.948
WT ∗ Path	6	315.889	52.648	22.037
WT ∗ Sample	6	166.335	27.723	11.604
Path ∗ Sample	4	38.050	9.513	3.982
Residuals	189	451.547	2.389	

interactions are important, the form of the WT ∗ Sample interaction is different. For all four waypoint times, increasing the amount of data tends to give smaller estimation errors. The rate of improvement is best when the waypoint time is specified to be 100 (the true value) and worst when the waypoint time is 120, which is furthest from the true value.

Now we describe the results based on the estimated waypoint location. For the Bayesian estimate, the relationship between the estimated waypoint location and WT is approximately linear, but with different slopes for each level of Sample. The slope is greatest with 75 observations and least with 50 observations. Also important in describing the estimated waypoint location, are the Path ∗ Sample interaction, and Var. For the Bayesian estimate, the Path ∗ WT interaction does not seem to be important. For the regression-based estimate, the relationship between estimated waypoint location and WT is linear, with different slopes for each level of Path (instead of for each level of Sample). This result is not surprising since the form of the estimate as shown in (8) and (9) is linear in waypoint time.

One other difference between the two estimators are the slopes of the relationship between estimated waypoint location and WT. For the Bayesian estimate, the average slopes for the x- and y-coordinates are 0.69 and 0.70. However, for the regression-based estimate, the corresponding average slopes are 0.94 and 1.01. As the Bayesian estimator can be thought of as a shrinkage estimator, the smaller slopes should be expected. Also, the slopes for the regression-based estimator are around 1, the average velocity in each direction, which is to be expected given the form of the estimator.

Finally, we describe the results based on the log determinant of the variance of the posterior distribution of the waypoint location, although obviously this is possible only for the Bayesian

Table 2. Analysis of Variance (ANOVA) table of estimation errors for regression-based waypoint-location estimate.

Factor	*Df*	Sum of Sq	Mean Sq	*F* Value
WT	3	8176.959	2725.653	3881.333
Path	2	2341.262	1170.631	1666.980
Sample	2	171.780	85.890	122.308
Var	1	0.383	0.383	0.545
Corr	2	0.578	0.289	0.411
WT ∗ Path	6	883.964	147.327	209.794
Path ∗ Sample	4	91.917	22.979	32.723
WT ∗ Sample	6	32.832	5.472	7.792
Residuals	189	132.725	0.702	

Table 3. Analysis of Variance (ANOVA) table for $\log|\Sigma_W|$.

Factor	*Df*	Sum of Sq	Mean Sq	*F* Value
Sample	2	58.889	29.445	427.803
WT	3	44.412	14.804	215.085
Var	1	4.008	4.008	58.230
Corr	2	1.515	0.758	11.006
Path	2	0.125	0.062	0.905
WT * Sample	6	5.446	0.908	13.189
WT * Corr	6	1.653	0.276	4.003
Residuals	193	13.284	0.069	

estimator. (The regression-based estimator has no valid measure of variability.) The precision of the Bayesian estimator, as measured by $\log|\Sigma_W|$, depends on four of the five possible factors, Sample, WT, Var, and Corr, as can be seen for the formal ANOVA table given in Table 3. Path does not seem to affect this measure. The two most important factors are WT and Sample. In addition, these two factors interact. The main trend is for the log determinant to decrease with increasing percentages of data (Sample) and to increase as WT increases. This should be expected, since the greater the difference between the assumed waypoint time and the last observation, the greater the opportunity for the paths to deviate from the trend at the end of the observed path. The effect of increasing the percentage of data sampled can be seen in Figures 3 and 4. The posterior distribution of the waypoint is much more spread out for 50% sampling than for 75% sampling. In addition, increasing Var tends to result in a small increase of the log determinant. Finally, Corr also has a small effect, as the high-correlation case tends to lead to a small decrease in the log determinant.

There is one other important observation that comes from this analysis. In all 216 cases, $\log|\Sigma_W|$ is smaller than the log determinant of the prior variance, implying in all cases that there is information in the data about the waypoint. However, in some cases, particularly with 50% sampling and a specified waypoint time of 120, the amount of information in the data about the waypoint is small.

6.6. Conclusions

The most important result from the simulation experiment is that the choice of estimation method is very important. We noted earlier that the Bayesian approach is more accurate in almost every case in the study. As can be seen in Figure 5, in the 21 cases (out of 216) where the regression-based estimate does better than the Bayesian estimate, the advantage is small and may be attributable to Monte Carlo sampling error (there are no obvious common features among these 21 cases). Moreover, for many of the cases where the Bayesian estimate dominates, its advantage is substantial. Intuitively, we could have expected this, since the Bayesian estimate uses more information (e.g., state equations, measurement-error variation); our study has quantified this intuition.

The next most important factor arising from the simulation experiment is waypoint time (WT). Misspecification of the WT may lead to large estimation errors for both estimators. This is expected and is due to the approximately linear relationship between the waypoint location estimates and WT.

Of the factors remaining, Sample has an effect on estimation accuracy d and, importantly, on estimation precision (as measured by $\log|\Sigma_W|$). As expected, the variance of the Bayesian

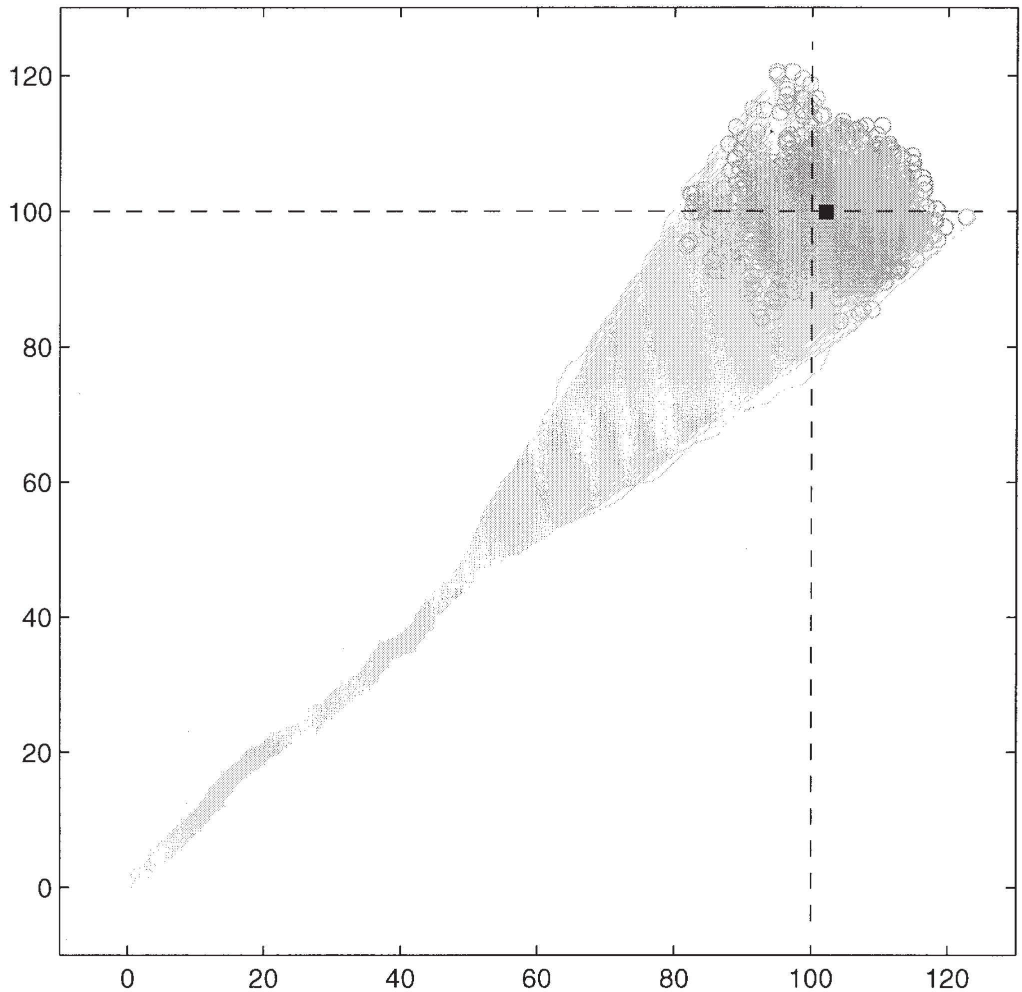

Figure 3. MCMC results for path 1, low variance, zero correlation, 50 observations, and specified waypoint time = 100. The light gray lines indicate the 2000 Monte Carlo realizations of the true path, the darker gray circles are the 2000 realizations of the waypoint locations, and the square is the Bayesian waypoint–location estimate. The dashed lines intersect at the true waypoint location (100, 100).

estimate decreases as more data are collected. WT also affects estimation precision, as smaller waypoint times lead to lower variability in the Bayesian estimates. The remaining factors (Path, Var, and Corr) in the simulation experiment were much less important.

There is clearly an investment of computing time needed to produce Bayesian estimates in this complex, nonlinear problem, where movement is determined by the start waypoint and an (unknown) end waypoint. If the time scale between waypoints is on the order of hours and a decision involving the object's destination is not very time-sensitive, the Bayesian approach yields both accurate and precise estimates. Consider the example where the waypoint is a mine-drop location. Time spent computing a better estimator and measures of its variability could well be repaid in terms of time spent searching for the enemy mines. However, this is not always the case; time may be of the essence for a C2 decision, where an estimate is needed

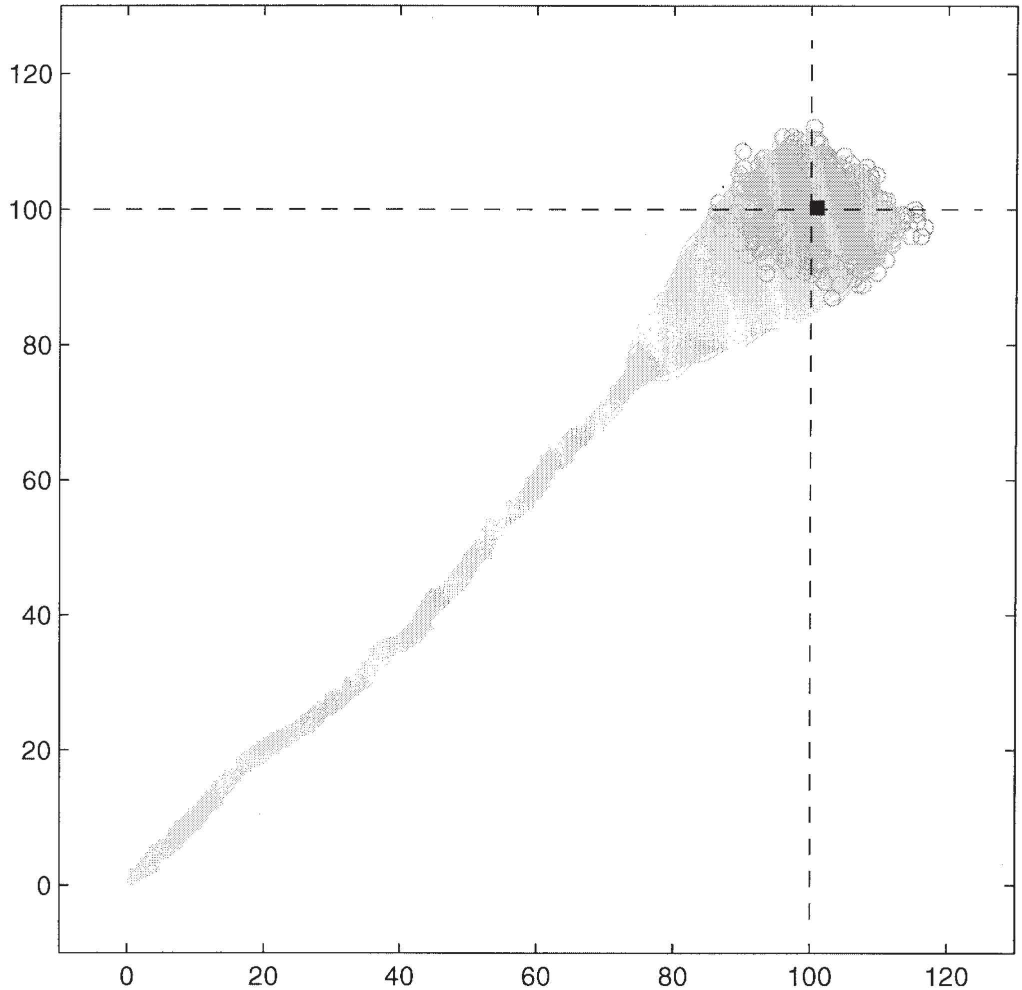

Figure 4. MCMC results for path 1, low variance, zero correlation, 75 observations, and specified waypoint time = 100. The light gray lines indicate the 2000 Monte Carlo realizations of the true path, the darker gray circles are the 2000 realizations of the waypoint locations, and the square is the Bayesian waypoint–location estimate. The dashed lines intersect at the true waypoint location (100, 100).

immediately or lives might be lost. The regression-based waypoint estimate is almost instantaneously computed and then would be the estimate of choice, by default. (However, it does not come with a valid measure of variability.) Because these two estimators are so different in terms of computing time, it is clear what the tradeoff is between accuracy and computing time. For comparisons between optimization algorithms, where accuracy of the estimator is a function of computing time, we refer the reader to the excellent article by Barr et al. [1].

7. DISCUSSION

There are several generalizations that can be made to this problem. First, we could consider additional objects in the battlespace. While this presents a complication to the overall problem

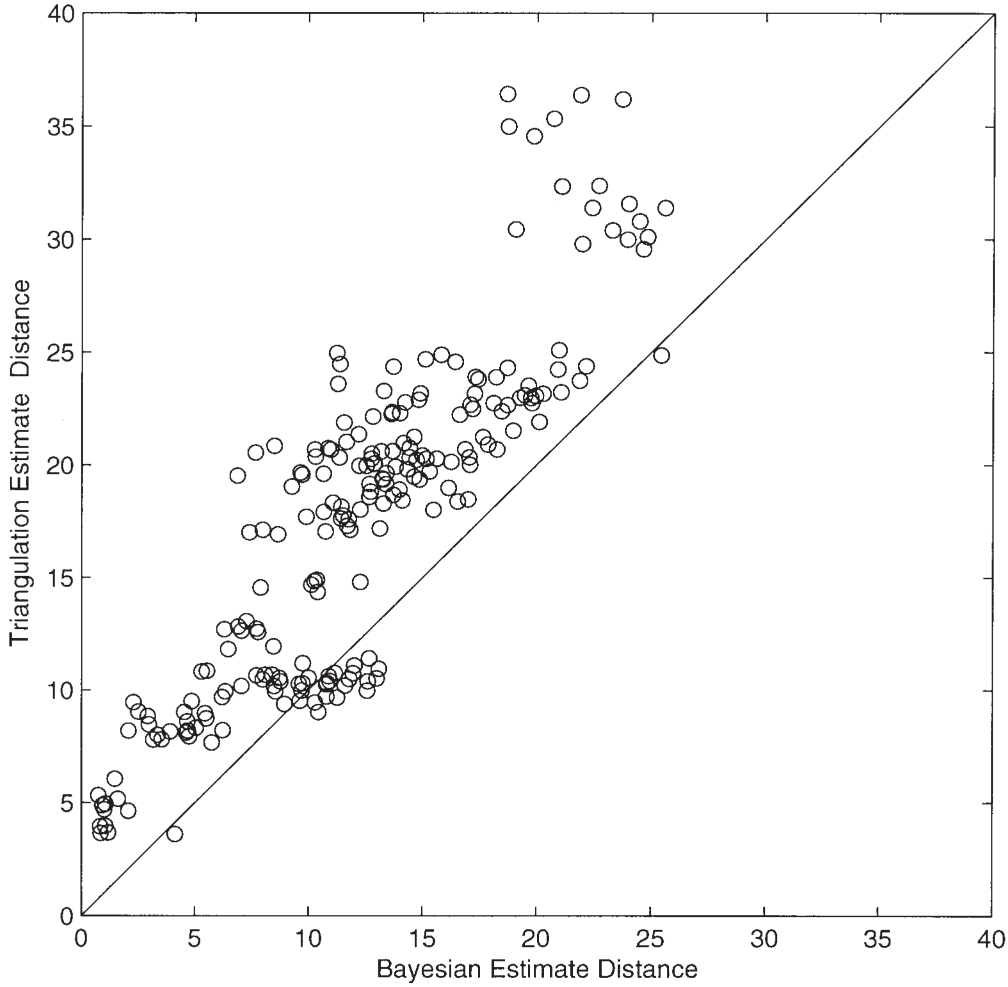

Figure 5. Comparison of $d(\hat{\mathbf{W}}_s, \mathbf{W}_s)$ for all 216 analysis conditions. Points lying above the diagonal line correspond to the Bayesian estimate being superior to regression-based estimate.

when several objects are headed towards the same waypoint, there is also an advantage in that there is additional data about the waypoint location. A second generalization would be to consider more complicated terrain. Rather than the flat terrain considered in this article, one might construct a terrain from piecewise planes, in a manner analogous to degree-0 splines. This is similar to the concept of triangulated irregular networks (TINs). In this more complicated battlespace, the impact of the terrain on movement, and thus on the state equations, must be considered. If the movement equations can be determined, the MCMC approach described in Section 4 can potentially be modified for this more complicated situation.

One issue that requires future study is estimation of the waypoint time. The simulation study focused on the distribution of the waypoint location at several prespecified waypoint times, which allowed for an examination of the sensitivity of the waypoint-location estimates to the specification of waypoint time. However, it does not directly address the estimation of the waypoint time or, more importantly, determination of its posterior distribution. We have already mentioned in Section 2.3 that the waypoint location might be known from intelligence sources; then the problem becomes Bayesian estimation of the waypoint given the waypoint location. This is a topic for future research.

In addition to the increased accuracy of the Bayesian estimator, it has other advantages over the regression-based estimator. It is accompanied by a natural measure of variability, namely the posterior variance of the waypoint location. Furthermore, the Monte Carlo scheme also gives estimates of the path taken by the object on its way to the waypoint (see Figs. 3 and 4). In contrast, no realistic path estimate is available for the regression-based procedure.

APPENDIX A

Using the assumptions given in Section 2.2 and from basic integration, the evolution of the state space can be represented as follows:

$$x_{k+1} = x_k + \int_k^{k+1} (s_k + (s_{k+1} - s_k)(t-k))\cos(\theta_k + (\theta_{k+1} - \theta_k)(t-k))\, dt. \tag{11}$$

Then, upon substituting $u = \theta_k + (\theta_{k+1} - \theta_k)(t-k)$, we obtain

$$\begin{aligned} x_{k+1} &= x_k + \frac{1}{\theta_{k+1} - \theta_k} \int_{\theta_k}^{\theta_{k+1}} \left(s_k + (s_{k+1} - s_k)\left(\frac{u - \theta_k}{\theta_{k+1} - \theta_k}\right)\right) \cos u\, du \\ &= x_k + \frac{1}{\theta_{k+1} - \theta_k} s_{k+1} \sin \theta_{k+1} - \frac{1}{\theta_{k+1} - \theta_k} s_k \sin \theta_k \\ &\qquad + \frac{1}{\theta_{k+1} - \theta_k} \frac{s_{k+1} - s_k}{\theta_{k+1} - \theta_k} (\cos \theta_{k+1} - \cos \theta_k). \end{aligned} \tag{12}$$

Likewise,

$$\begin{aligned} y_{k+1} &= y_k + \int_k^{k+1} (s_k + (s_{k+1} - s_k)(t-k))\sin(\theta_k + (\theta_{k+1} - \theta_k)(t-k))\, dt \\ &= y_k - \frac{1}{\theta_{k+1} - \theta_k} s_{k+1} \cos \theta_{k+1} + \frac{1}{\theta_{k+1} - \theta_k} s_k \cos \theta_k \\ &\qquad + \frac{1}{\theta_{k+1} - \theta_k} \frac{s_{k+1} - s_k}{\theta_{k+1} - \theta_k} (\sin \theta_{k+1} - \sin \theta_k). \end{aligned} \tag{13}$$

APPENDIX B

Let $\pi(x)$ be the density of the desired target (Gibbs) sampling distribution and suppose ℓ points, $x^{(1)}, \ldots, x^{(\ell)}$, have been sampled. Assume that the target distribution is difficult to sample from, but that the density function $\pi(\cdot)$ is easy to calculate, up to a constant of proportionality. Further, let $f(y|x)$ be a probability density function describing the proposal distribution for sampling a new state y given the current state is x. Then the *Metropolis-Hastings algorithm* is given as follows:

- Simulate y from the proposal distribution $f(y|x^{(\ell)})$.
- Simulate $U \sim$ Uniform[0, 1] and define

$$x^{(\ell+1)} = \begin{cases} y & \text{if } U \le r(x^{(\ell)}, y), \\ x^{(\ell)} & \text{otherwise,} \end{cases}$$

where

$$r(x, y) \equiv \min\left\{1, \frac{\pi(y) f(x|y)}{\pi(x) f(y|x)}\right\},$$

is known as the Hastings ratio. For more details, see Gelman et al. [4], Section 11.2.

ACKNOWLEDGMENTS

This research was supported the Office of Naval Research under Grants N00014-99-1-0214 and N00014-02-1-0052. The authors would like to express their appreciation to Gardar Johannesson and John Kornak for a number of helpful suggestions, and to two referees and the Associate Editor for their constructive comments that led to a substantial revision of the original submission.

REFERENCES

[1] R.S. Barr, B.L. Golden, J.P. Kelly, M.G.C. Resende, and W.R. Stewart, Designing and reporting on computational experiments with heuristic methods, J Heuristics 1 (1995), 9–32.
[2] L.M. Berliner, "Hierarchical Bayesian time series models," Maximum entropy and Bayesian methods, K. Hanson and R. Silver (Editors), Kluwer, Dordrecht, 1996, pp. 15–22.
[3] B.P. Carlin and T.A. Louis, Bayes and empirical Bayes method for data analysis, 2nd edition, CRC/Chapman and Hall, Boca Raton, FL, 2003.
[4] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin, Bayesian data analysis, Chapman and Hall, New York, 1995.
[5] M.E. Irwin, N. Cressie, and G. Johnannesson, Spatial-temporal nonlinear filtering based on hierarchical statistical models, Test 11 (2002), 249–302.
[6] E.M. Paté-Cornell and P.S. Fischbeck, Probabilistic interpretation of command and control signals: Bayesian updating of the probability of nuclear attack, Reliab Eng Syst Safety 47 (1995), 27–36.
[7] E.T. Sherrill and D.R. Barr, Exploring a relationship between tactical intelligence and battle results, Mil Oper Res 2 (1996), 17–33.
[8] R. van der Merwe, A. Doucet, N. de Freitas, and E. Wan, The unscented particle filter, Technical report CUED/F-INFENG/TR 380, Department of Engineering, Cambridge University, 2000.
[9] D. Wendt, N. Cressie, and G. Johannesson, "A spatial-temporal statistical approach to command and control," Battlespace digitization and network-centric warfare, R. Suresh (Editor), SPIE Proc 439, SPIE, Bellingham, WA, 2001, pp. 232–243.