

Autonomous Decision Making with Uncertainty for an Urban Intelligence, Surveillance and Reconnaissance (ISR) Scenario

Raymond W. Holsapple * and Phillip R. Chandler †

Air Force Research Laboratory, Wright-Patterson AFB, Ohio, 45433, USA

John J. Baker ‡ and Anouck R. Girard §

University of Michigan, Ann Arbor, Michigan, 48109, USA

Meir Pachter ¶

Air Force Institute of Technology, Wright-Patterson AFB, Ohio, 45433, USA

In this paper, we consider an urban ISR scenario where a human operator is tasked to provide feedback regarding the nature of some objects of interest. The feedback is relayed to the stochastic controller of an unmanned aerial vehicle (UAV), which must determine an appropriate mission plan. A small (unmanned) aerial vehicle (SAV) loiters at a high altitude where it may survey a large territory. An operator decides which objects in the SAV's field of view are of interest and which are not. Then a team of micro (unmanned) aerial vehicles (MAVs) is assigned individual tours to inspect the objects of interest at a low altitude. As a MAV flies over an object of interest, the operator must decide if the object has a feature that uniquely distinguishes it as a target. The key parameters are the operator's response and the time taken for the operator to respond. The controller uses these parameters to compute the expected information gain of a revisit. In previous studies automatic target recognition (ATR) was used for making some decisions in the SAV and the MAVs. This paper investigates the use of human feedback alone for target recognition. Different methods for calculating expected information gain are examined and compared. In addition, results from a flight test of this controller are presented.

Nomenclature

T	Target truth status
V	Visibility truth status for a target-distinguishing feature on a specific flyover
R	Response from operator for a target-distinguishing feature being present on a specific flyover
θ	Target-distinguishing feature visibility arc angle

Subscripts

1	First flyover of an object of interest
2	Second flyover of an object of interest
t	Boolean: true
f	Boolean: false

*Research Mathematician, Control Sciences Center of Excellence, Wright-Patterson AFB, OH 45433, AIAA Member.

†Technical Advisor, Control Sciences Center of Excellence, Wright-Patterson AFB, OH 45433, AIAA Senior Member.

‡Research Assistant, Department of Aerospace Engineering, Ann Arbor, MI 48109.

§Assistant Professor, Department of Aerospace Engineering, Ann Arbor, MI 48109, AIAA Member.

¶Professor, Department of Electrical and Computer Engineering, Wright-Patterson AFB, OH 45433, AIAA Associate Fellow.

I. Introduction

A. COUNTER Scenario

THE decision algorithms discussed in this paper are tailored for the COUNTER¹ scenario. COUNTER is an acronym for Cooperative Operations in Urban Terrain, an Air Force program using a team of UAVs to investigate task assignment and path planning algorithms for use in ISR missions in urban areas.

COUNTER uses a team of UAVs—one SAV and four MAVs. The SAV loiters over the urban area at 1000-1500 feet above ground level (AGL), while an operator views the live video feed from the SAV for objects of interest. We assume that the objects remain stationary. After an operator selects a collection of objects to view more closely, a task assignment algorithm assigns a tour to each MAV that is to be launched. The MAVs fly at a much lower altitude (50-150 feet AGL) allowing them to inspect the objects of interest close up and at an acute angle, which may permit them to see into vehicles and under tarpaulins and camouflage nets. Like the SAV, each MAV is equipped with front and side facing video cameras. This video feed is relayed back to a ground control station where an operator attempts to classify the objects in real time as the MAVs inspect the collection of objects assigned to them.

The operator is not asked to give a response whether or not a particular object of interest is a target or a non-target based on his or her inspection of the video. Instead, the operator is asked whether or not he or she has seen a distinguishing feature that has been described to him or her prior to the mission. The operator may even have a sample picture of such a feature to refer to during the mission. The assumption about this feature is that it uniquely separates targets from non-targets.

B. Stochastic Controller

It was anticipated that a human operator would be overwhelmed if expected to manage MAV tour reassignments while simultaneously attempting to detect object features in video. Therefore, a stochastic controller was developed to decide when a revisit for additional information is needed.² The key feature of this approach is the inclusion of an operator error model. Stochastic dynamic programming is used to solve this decision making with uncertainty problem. The dynamic program state is the amount of fuel that is allocated for revisiting objects of interest, hereafter referred to as reserve.

In addition to the reserve, the stochastic controller also makes use of the operator's response, the amount of time the operator took to make the response (operator delay), and the number of remaining objects in the MAV's tour. Given the operator's response, the controller performs an information gain analysis where it computes an expected reward for performing a revisit using *a priori* probabilities that were determined experimentally. These probabilities characterize the target density and operator's decision behavior. Given the operator's response, operator delay, reserve, and remaining objects, the controller will decide whether a MAV should revisit an object and modify the MAV's flight plan accordingly.

Revisits are often useful because they can provide additional information regarding an object of interest at a different approach angle. For example, a feature may only be visible from the rear aspect of the object, so if the MAV approaches from the front aspect only, the operator will never see the definitive feature. In this case there may (or may not) be sufficient information gain to perform a revisit, depending on the values of the inputs to the controller. In the case of extremely low target density, there might be sufficient information gain to perform a revisit even in cases where the operator says that he or she has seen the distinguishing feature on the initial inspection. This is due to the uncertainty in the operator's responses, which will be discussed later.

Here we note that the reserve set aside for revisits is finite. As previously mentioned, this reserve is the state of the dynamic program that is used to solve this optimization problem. The solution of the dynamic program generates a matrix of cost thresholds that will be used in the decision process. At the moment an operator gives a response (which coincides with the moment of decision by the controller), the expected cost of a revisit is compared to the cost threshold which was computed in the information gain analysis. At this point the control decision is simply a table look up. If the expected cost is less than the cost threshold, the MAV will revisit the object of interest.

C. Motivation

In previous publications²⁻⁶ a probabilistic method was developed to determine the reward values that are used in the cost function of the dynamic program. The probabilities were developed with the assumption

that the MAV had an ATR device used for classifying objects of interest. The MAV would only defer to an operator for classification if the ATR encountered an ambiguous feature.

Optical feature recognition would be problematic in application because of additional payload constraints, communication unreliability, and highly variable lighting conditions. Additionally, and perhaps more practically, the MAVs used for COUNTER are simply not equipped with such a device. This was the motivation for a system that instead of employing an ATR, relies solely on an operator for feature recognition.

D. Original Contributions

In this paper, three different expected reward functions will be considered. These functions will rely on the operator’s response from the flyover of an object of interest instead of utilizing an ATR. The number of a posteriori probabilities needed for the reward functions will be developed. Then an analysis will be performed comparing the reward methods against a benchmark and each other. Performance of each method will be discussed and a recommendation for future work will be made.

II. Formulation

A. Definition of Terms

In this system there are three main events, and they can be treated as boolean operators: Feature Visibility, Operator Response, and Target Truth Status (whether or not an object of interest is an actual target). These are considered boolean because only absolutes are considered. For example, the feature is either visible or not visible, the operator indicates feature or no feature, and an object is either a target or it is not. Each MAV visit can be thought of as some combination of these events. The first and second visit of a MAV will be treated as independent events, so the subscripts denoting the visit are only important when probabilities involve events from both visits.

B. A Priori Probabilities

The probability that an object of interest is a target is assumed *a priori*:

$$\begin{aligned} P(T_t) &= p, \\ P(T_f) &= 1 - p. \end{aligned} \tag{1} \tag{2}$$

In some scenarios p can be very low; it could be one out of thousands in extremely cluttered urban areas. For the purposes of this paper, we chose values of p in the interval $[0.1, 0.2]$. If we were to let p be extremely small, we would have to load the simulation with hundreds or thousands of objects of interest each time the simulation was executed. This would lead to an extraordinarily large amount of time spent on running simulations. Moreover, this value of p is the *a priori* probability used in the COUNTER flight tests. An operator confusion matrix developed in² is a way of depicting the stochastic behavior of a human operator given a collection of probabilistic events. In previous efforts,²⁻⁶ the probabilistic event corresponding to any given operator response was simply the target truth status. This truth status is of course unknown to the operator but has a probability distribution described by equations (1) and (2).

Table 1. Operator Confusion Matrix

	R_t	R_f
$P(R V_t \cap T_t)$	P_D	$1 - P_D$
$P(R V_t \cap T_f)$	undefined	undefined
$P(R V_f \cap T_t)$	P_{FA}	$1 - P_{FA}$
$P(R V_f \cap T_f)$	P_{FA}	$1 - P_{FA}$

The operator confusion matrix in this paper is different from the previous version because it accounts for an additional stochastic event, the feature visibility. Note that a feature cannot be visible on an object of

interest that is not an actual target, thus in Table 1, $P(R|V_t \cap T_f)$ is undefined by nature. A design choice was made that the probability that an operator responds that an object of interest has the feature, when no feature is actually visible and the object is in fact a target, would simply be the probability of a false alarm. The issue is that even if the operator's implication (the object is a target) is technically correct, it is based on no visual evidence and should be treated as a false alarm, which is how it is modeled in this paper. The operator confusion matrix is shown in Table 1.

Here we should note a couple of things. First, P_D (probability of detection) and P_{FA} (probability of false alarm) are conditional probabilities of the operator's response given the target truth status and the feature visibility status. Secondly, it is assumed that P_D and P_{FA} are affected by the operator's workload. Nominally we suggest that as an operator's workload increases, the probability of detection should decrease while the probability of false alarm increases.

Next consider the possible feature visibility outcomes from a MAV flying over an object of interest. Each target is modeled as having its distinguishing feature visible only when the target is approached from a heading that falls within an assumed arc angle, θ . The range of visibility divided by the total range of angles the object can be viewed from is the conditional probability that a feature is visible given that it is a target. Table 2 lists this set of *a priori* conditional probabilities.

Table 2. Visibility Given Target Truth Status

	V_t	V_f
$P(V T_t)$	$\theta/2\pi$	$1 - \theta/2\pi$
$P(V T_f)$	0	1

In the case of two visits, the system is modeled such that the MAVs perform their second visit from the opposite angle of approach. According to the model, if a feature is visible on the first pass, it will not be visible on the second pass and vice versa. Given these constraints, the conditional probabilities of feature visibility for two visits given the target truth status may be inferred. These conditional probabilities are given in Table 3.

Table 3. Visibility Given Target Truth Status

	V_{1t}, V_{2t}	V_{1t}, V_{2f}	V_{1f}, V_{2t}	V_{1f}, V_{2f}
$P(V_1 \cap V_2 T_t)$	0	$\theta/2\pi$	$\theta/2\pi$	$1 - \theta/\pi$
$P(V_1 \cap V_2 T_f)$	0	0	0	1

C. Reward Multiplier Probabilities

The motivation of the following exercise is to determine the probability of the operator's decision and target feature visibility on a second visit, given the decision and visibility from the first visit. These probabilities will be used as gains applied against reward values so that they are weighted according to the probability that they will occur. To do this, the probabilities will be broken down into their constituent *a priori* sub-probability combinations. Before describing how that is done, we first note that we seek the following:

$$\begin{aligned}
 P(R_2 \cap V_2|R_1 \cap V_1) &= \frac{P((R_2 \cap V_2) \cap (R_1 \cap V_1))}{P(R_1 \cap V_1)} \\
 &= \frac{P(R_2 \cap V_2 \cap R_1 \cap V_1 \cap T_t) + P(R_2 \cap V_2 \cap R_1 \cap V_1 \cap T_f)}{P(R_1 \cap V_1 \cap T_t) + P(R_1 \cap V_1 \cap T_f)}. \tag{3}
 \end{aligned}$$

The terms in the denominator can be resolved into their *a priori* constituent parts as follows:

$$\begin{aligned} P(R_1 \cap V_1 \cap T) &= P(R_1|V_1 \cap T)P(V_1 \cap T) \\ &= P(R_1|V_1 \cap T)P(V_1|T)P(T) \end{aligned} \quad (4)$$

$$\equiv \tilde{P}(T). \quad (5)$$

The terms in the numerator may also be resolved into *a priori* constituents. To do this we will begin with a definition.

Definition 1 Two events E_1 and E_2 are conditionally independent of event E_3 if and only if

$$P(E_1 \cap E_2|E_3) = P(E_1|E_3)P(E_2|E_3), \quad (6)$$

or equivalently

$$P(E_1|E_2 \cap E_3) = P(E_1|E_3). \quad (7)$$

While breaking the terms in the numerator into their constituent parts, we must assume conditional independence several times. In the equations below, conditional independence is assumed as we proceed from equation (8) to equation (9) and from equation (10) to equation (11). This assumption is intuitive, as it makes sense to assume that the operator's response and feature visibility pairs from the first and second visits should be conditionally independent of each other given the target truth status. The terms in the numerator are broken down as follows:

$$P(R_2 \cap V_2 \cap R_1 \cap V_1 \cap T) = P((R_2 \cap V_2) \cap (R_1 \cap V_1)|T)P(T) \quad (8)$$

$$= P(R_2 \cap V_2|T)P(R_1 \cap V_1|T)P(T) \quad (9)$$

$$= \frac{P(R_2 \cap V_2 \cap T)}{P(T)} \frac{P(R_1 \cap V_1 \cap T)}{P(T)} P(T)$$

$$= \frac{P(R_2 \cap V_2 \cap T)P(R_1 \cap V_1 \cap T)}{P(T)}$$

$$= \frac{P(R_2|V_2 \cap T)P(V_2 \cap T)P(R_1|V_1 \cap T)P(V_1 \cap T)}{P(T)}$$

$$= P(R_2|V_2 \cap T)P(R_1|V_1 \cap T) \frac{P(V_2|T)P(T)}{P(T)} P(V_1|T)P(T)$$

$$= P(R_1|V_1 \cap T)P(R_2|V_2 \cap T)P(V_1|T)P(V_2|T)P(T) \quad (10)$$

$$= P(R_1|V_1 \cap T)P(R_2|V_2 \cap T)P(V_1 \cap V_2|T)P(T) \quad (11)$$

$$\equiv \hat{P}(T). \quad (12)$$

Finally, we may use equations (4) and (11) to assemble the *a priori* constituent forms of the numerator and denominator in equation (3). For brevity we will write (3) using the equivalent definitions given by equations (5) and (12),

$$P(R_2 \cap V_2|R_1 \cap V_1) = \frac{\hat{P}(T_t) + \hat{P}(T_f)}{\tilde{P}(T_t) + \tilde{P}(T_f)}. \quad (13)$$

D. Reward Probabilities

The two conditional probabilities used to compute the reward values must also be determined. These two probabilities are $P(T|R_1 \cap V_1)$ and $P(T|R_1 \cap V_1 \cap R_2 \cap V_2)$, and they are decomposed into constituents in the equations below. Using equations (4) and (5), we have

$$\begin{aligned} P(T|R_1 \cap V_1) &= \frac{P(T \cap R_1 \cap V_1)}{P(R_1 \cap V_1)} \\ &= \frac{P(R_1 \cap V_1 \cap T)}{P(R_1 \cap V_1 \cap T_t) + P(R_1 \cap V_1 \cap T_f)} \\ &= \frac{\tilde{P}(T)}{\tilde{P}(T_t) + \tilde{P}(T_f)}. \end{aligned} \quad (14)$$

Furthermore, using equations (11) and (12), the more complicated of the two conditional probabilities becomes

$$\begin{aligned}
P(T|R_1 \cap V_1 \cap R_2 \cap V_2) &= \frac{P(T \cap R_1 \cap V_1 \cap R_2 \cap V_2)}{P(R_1 \cap V_1 \cap R_2 \cap V_2)} \\
&= \frac{P(R_2 \cap V_2 \cap R_1 \cap V_1 \cap T)}{P(R_2 \cap V_2 \cap R_1 \cap V_1 \cap T_t) + P(R_2 \cap V_2 \cap R_1 \cap V_1 \cap T_f)} \\
&= \frac{\hat{P}(T)}{\hat{P}(T_t) + \hat{P}(T_f)}. \tag{15}
\end{aligned}$$

Since each event is represented by a boolean value, the number of equations needed to describe the system is simply 2^n , where n is the number of boolean events involved. For example, $n = 5$ in equation (15), so there are thirty-two equations like equation (15).

E. Reward Functions

Two information theory reward functions from a previous effort² and an additional method, where discrete reward values are assigned, will be considered and evaluated. The range of values from the rewards is essentially arbitrary but provides a basis for comparison of possible outcomes. These rewards will then be scaled respectively by equation (13) from the previous section. For the purpose of brevity while describing the three methods, let $A = R_1 \cap V_1$ and $B = R_2 \cap V_2$.

Method 1 We begin by defining some conditional probabilities:

$$P_{11} = P(T_t|A), \tag{16}$$

$$P_{12} = P(T_f|A), \tag{17}$$

$$P_{13} = P(T_t|A \cap B), \tag{18}$$

$$P_{14} = P(T_f|A \cap B). \tag{19}$$

Then the reward value using Method 1 can be expressed using equations (16) - (19) and is given by the following:

$$R_1 = \log \left(\frac{P_{13}}{P_{14}} + \frac{P_{14}}{P_{13}} \right) - \log \left(\frac{P_{11}}{P_{12}} + \frac{P_{12}}{P_{11}} \right). \tag{20}$$

Method 2 We begin by defining some conditional probabilities:

$$P_{21} = P(T_t \cap A), \tag{21}$$

$$P_{22} = P(T_f \cap A), \tag{22}$$

$$P_{23} = P(T_t \cap A \cap B), \tag{23}$$

$$P_{24} = P(T_f \cap A \cap B). \tag{24}$$

Then the reward value using Method 2 can be expressed using equations (1), (2), (16) - (19), (21) - (24) and is given by the following:

$$\begin{aligned}
R_2 &= \left(P_{23} \log \left(\frac{P_{13}}{p} \right) + P_{24} \log \left(\frac{P_{14}}{1-p} \right) \right) - \\
&\quad \left(P_{21} \log \left(\frac{P_{11}}{p} \right) + P_{22} \log \left(\frac{P_{12}}{1-p} \right) \right). \tag{25}
\end{aligned}$$

Method 3 For method three, discrete values were chosen for the sixteen combinations of operator response and feature visibility for both visits. A value of zero was assigned if the outcome was impossible, such as the feature being visible on both passes. A small reward was given if the operator was incorrect on both passes but the situation was possible, or when the operator was correct on the first visit but incorrect on the second visit. Moderate rewards were assigned for the operator being incorrect on the first pass, but correct on the second. The largest reward was given when the operator was correct on both visits. One benefit of this method is that there is never a negative reward value, which is possible with information theory and causes saturation within the revisit threshold function which is discussed in the following subsection.

Benchmark A comprehensive study using Monte Carlo simulations was done to determine the mean operator response delay. A time threshold slightly greater than the mean delay was chosen so that it would envelope a majority of the operator delay times. Thus, if the operator delay was less than the threshold, the UAV would perform a revisit.

F. Threshold Surface Plots

Analyzing the surfaces provided by the threshold function provides a preliminary indication of how the system will respond. The threshold surface is determined by the operator’s response, the amount of remaining reserve, the operator delay, the number of objects remaining to make a decision about and the expected reward.⁴⁻⁶ Essentially, the system determines an expected cost for revisit, and if that expected cost is less than the corresponding threshold value from the threshold surface, the MAV will perform a revisit. Having a deep intuitive understanding of how the shape of the threshold surface impacts the response of the system is useful but not necessary. In this paper, we are ultimately concerned with any saturation that occurs along the threshold value axis. Large amounts of saturation indicate a bias towards a certain stochastic controller decision.

For Method 1, Figure 1 indicates that the case where the operator responds with true, results in a very saturated threshold surface. This means that when the operator responds that they see a feature, the stochastic controller will virtually always perform a revisit. Whereas the case where the operator responds false, seen in Figure 2, is not as predictable.

For Method 2, Figure 3 indicates that if the operator responds with true, then the stochastic controller will most likely not make a revisit, whereas if the operator responds false, the stochastic controller will most likely perform a revisit as shown in Figure 4.

For Method 3, there is not enough saturation in Figures 5 and 6 to definitively say what the stochastic controller will decide to do given the operator’s response.

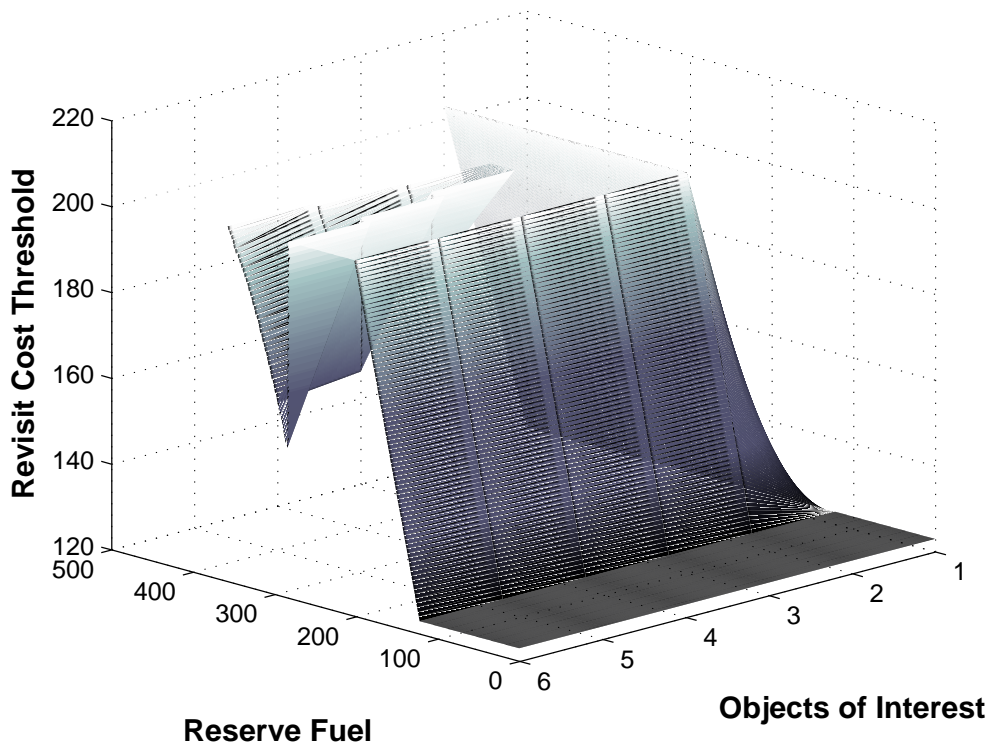


Figure 1. Method 1 Threshold Surface Plot, Operator Response is True

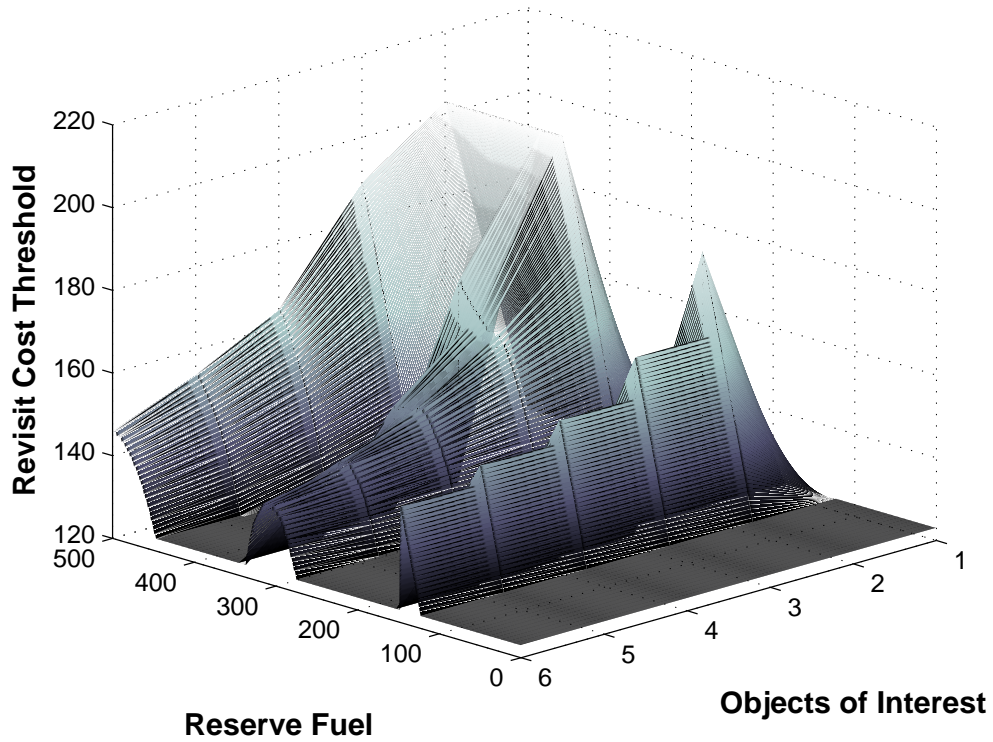


Figure 2. Method 1 Threshold Surface Plot, Operator Response is False

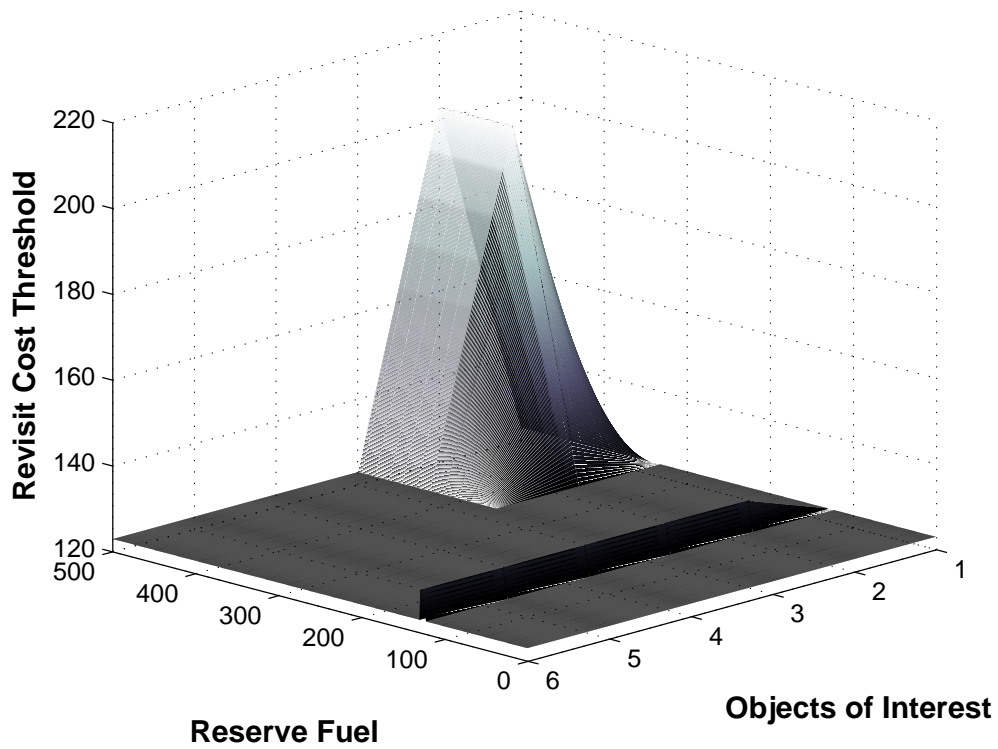


Figure 3. Method 2 Threshold Surface Plot, Operator Response is True

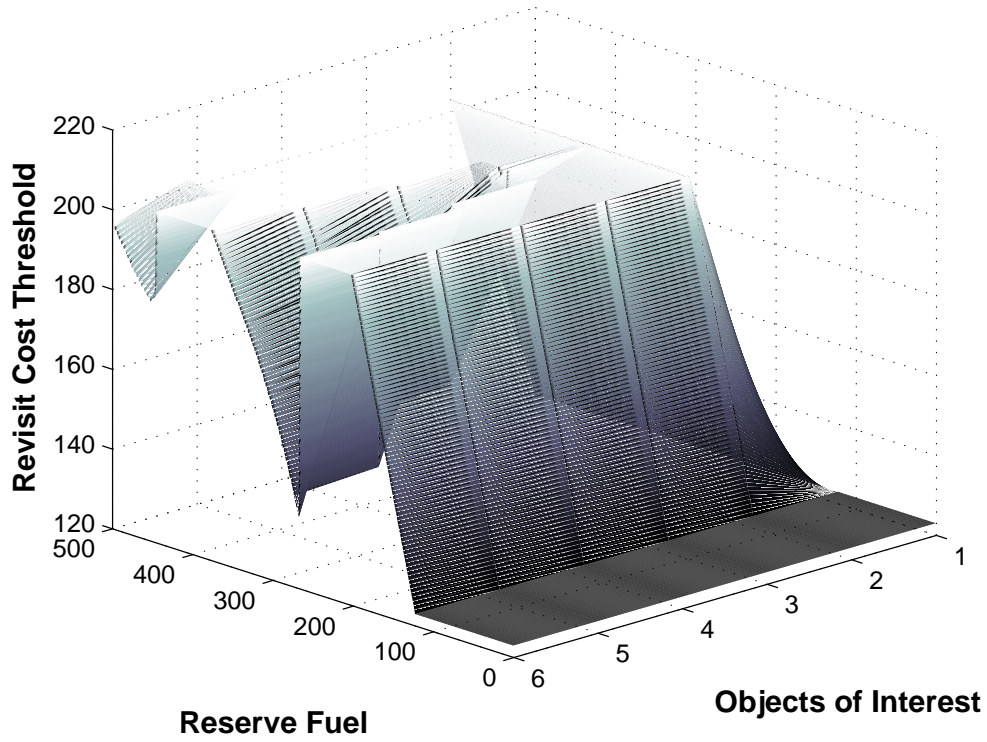


Figure 4. Method 2 Threshold Surface Plot, Operator Response is False

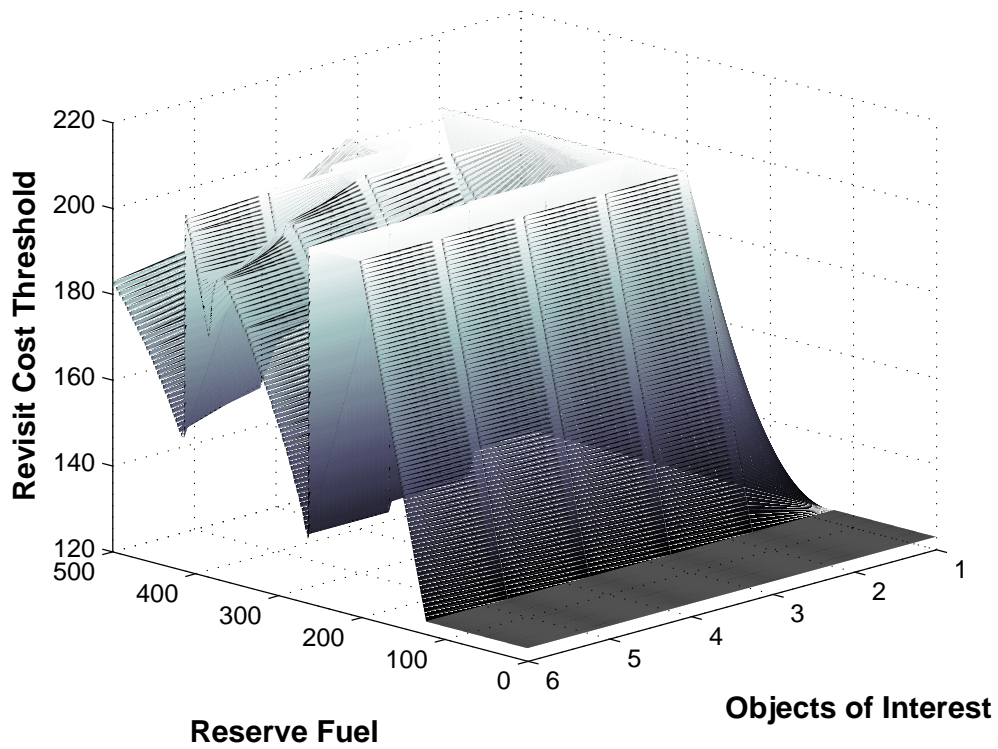


Figure 5. Method 3 Threshold Surface Plot, Operator Response is True

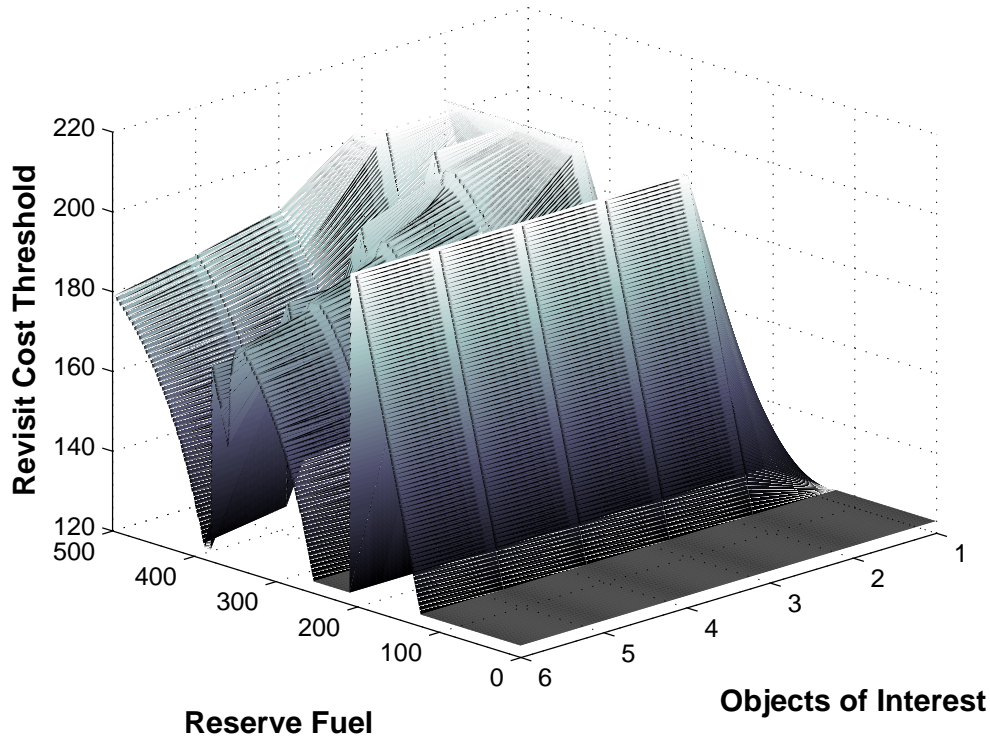


Figure 6. Method 3 Threshold Surface Plot, Operator Response is False

III. Simulation and Results

Simulations were conducted to test the various reward methods. Each method underwent 100 trials, each trial being 1200 simulation seconds long, with twenty objects of interest and four UAVs on tour. The location, orientation, and target truth status of the objects of interest were randomized for each trial. A log was kept of the stochastic controller’s decisions throughout all of the trials, the data from which acts as a basis for comparison between the various reward methods.

To compare the different methods, a rating system was devised. For all cases where the stochastic controller had a UAV revisit an object of interest, it added a point if either of the following occurred:

- the operator response was true on one visit and false on the other visit and the target truth status was true,
- the operator response was false on both visits and the target truth status was false.

These are the best case scenarios where everything the operator indicates coincides with reality. The points tallied for the case where target truth status is true and the case where it was false was kept separate because they have a different probability of occurrence. The overall point tally for each case is then divided by the probability of that case to determine a normalized point system. The points for each target truth status case can then be added together to determine an overall score for that method. Table 4 summarizes the results.

From Table 4, it can be seen that Method 2 outperforms the other methods for the case where the target truth status is false. This behavior likely results from the fact that Method 2 has a strong preference, due to threshold function saturation, to opt for a revisit when the operator responds ‘did not see feature’, in combination with the fact that an operator will tend to give this particular response on both visits more often than not simply due to target density. Method 3 outperforms the other methods for the case where the target truth status is true, and its overall score indicates that it is the best of the three methods. This could be due to a lack of significant saturation in both threshold functions seen in Figures 5 and 6. Although it

Table 4. Simulation Results

Simulation target density $\rightarrow p = 0.1$	Benchmark	Method 1	Method 2	Method 3
T_t mean score (S_t)	0.190	0.610	0.500	0.760
T_t standard deviation	0.419	0.803	0.674	0.842
T_f mean score (S_f)	0.410	0.070	2.080	0.520
T_f standard deviation	0.911	0.432	2.977	1.453
Adjusted S_t ($\tilde{S}_t = S_t/p$)	1.900	6.100	5.000	7.600
Adjusted S_f ($\tilde{S}_f = S_f/(1-p)$)	0.456	0.078	2.311	0.578
Total Score ($S = \tilde{S}_t + \tilde{S}_f$)	2.356	6.178	7.311	8.178

has no strong preference for either response, it seems to perform well on average for both, whereas Methods 1 and 2 tend to favor a particular response.

IV. COUNTER Flight Tests

In October 2007, the COUNTER team conducted a week-long flight test in 29 Palms, CA at the Marine Corp Air Ground Combat Center. The purpose of the flight test was to demonstrate various cooperative control technologies for a team of UAVs. The COUNTER team had conducted more than five successful flight tests prior to this one, but this was the first flight test where the main purpose was to test the stochastic controller outside of a simulation environment.

There were no flight tests designed to experimentally examine the performance of the stochastic controller using the various reward functions presented in the previous sections. All of the results from this paper are the result of simulations, a majority of which took place after the COUNTER flight test in 29 Palms. The main purpose of the test plan for this flight test was to experimentally investigate the original design of the stochastic controller.²⁻⁶

A. Test Plan

The test plan consisted of eight tests in three main categories. The first main goal of the tests was to examine the linear model of the expected revisit cost. Three tests were designed for this goal, one flown along a straight path and two flown along curved paths. The second main goal of the tests was to examine the decision behavior of the stochastic controller. Two tests were designed for this goal. The third main goal of the tests was to examine the operator model under various workloads. Three tests were also designed for this goal, one for each of light, medium and heavy workloads.

1. Linear Revisit Cost Model

The first test was designed to examine the theoretical model behind the design of the stochastic controller. The goal was to validate the linear model of the expected revisit cost, while the UAVs flew along a straight path. This is exactly how the theoretical model is addressed in previous publications.^{2-4,6} Data on the actual reserve used to revisit targets was collected for a range of operator delay times to make the comparisons to the model. The model being tested can be simply stated. Let τ_i be the random operator delay that is realized at the moment the operator gives his or her response to the stochastic controller about object i . Let τ_0 be the fixed delay associated with the time it takes for the MAV to turn around twice. Then the expected cost of a revisit, J , is computed as follows:

$$J(\tau_i) = 2\tau_i + \tau_0. \quad (26)$$

The second and third tests were also designed to examine the linear cost model given by equation (27). However, these tests were not aimed at testing how the theoretical model performed in a real flight test.

Instead, the purpose was to examine how the linear model would perform when the objects assigned for inspection do not lie in a straight line. This situation is the more realistic one for actual urban ISR missions. The only difference between tests two and three is the location of the objects of interest. The positions are random in both cases of course, but the objects were to be re-arranged for test three. The same data were to be collected for these tests as for test one.

2. Stochastic Controller Decision Behavior

The fourth and fifth tests were designed to investigate the qualitative behavior of the stochastic controller. No metric can accurately describe the intent of these two tests. Essentially, these tests were designed to determine the situations (values of operator delay, reserve, operator response, number of targets to visit) in which the controller would choose to revisit an object of interest. For instance, in situations where a MAV is visiting an object early in its tour (first or second) and the remaining reserve is high, what does the controller decide to do when the operator responds ‘saw feature’ quickly? What does it decide to do if in the same situation the operator responds ‘did not see feature’ slowly?

This qualitative study of the controller’s behavior could then be used by the designers of the stochastic controller to gain some insight into the decision behavior of the controller. It could even be used to compare the decisions of the controller to decisions that a human might make in the same situation without the aid of the controller. Data were to be collected for many permutations of all the possible situations. The initial reserve for test four was set so that only one revisit would be possible. The only differences between the two tests was that in test five the objects of interest were to be re-arranged and the initial reserve was set so that two revisits would be taken.

3. Operator Model

The sixth, seventh and eighth tests were designed to validate the operator model for low, medium and high workloads. We defined a low workload to mean that an operator was responsible for viewing video from two MAVs during the mission. Similarly, medium workload was three MAVs, and high workload was four MAVs. The number of objects of interest per MAV tour was designed to be five in all three tests. In addition, the target density in each test was 0.2. For example, medium workload meant the operator was responsible for three MAVs that would be inspecting 15 objects of interest with three true targets among them.

There were several key parameters these tests were designed to investigate. Data were to be collected to determine the operator’s missed detection rate and false alarm rate, which are the parameters that define the operator confusion matrix. Data were also to be collected to approximate the operator delay distribution. This distribution was to be compared to the distribution of the operator model. An approximation of the operator model’s distribution is seen in Figure 7. This was determined by collecting operator delay data from a Monte Carlo simulation using the operator model. The best fit of the data is a $\Gamma(1.4, 7.5)$ distribution.

B. Results

The results of the flight tests were not as complete as desired. We were not able to complete any of the flight tests as written in the test cards. No data were collected for tests three, five and eight. The amount of data collected for tests two, six and seven was not statistically significant; in some cases literally only a couple of data points were collected.

The reason for the low data collection is twofold. This was the first time the stochastic controller was fielded in a flight test, and there are always many unforeseen issues that need to be addressed and parameters that need to be adjusted to get a system to perform close to how it does in simulation. Secondly, the platform we were using for the MAVs proved to be more unreliable than expected. While testing the stochastic controller, roughly 29% of all MAV launches yielded any useful data for the test card being run. Moreover, some of those 29% yielded only partial data, not full data, for the particular test being performed. When these two issues are combined with other test cards which were also being executed in the four days of the flight test, one can see that the amount of data that could be collected on a first test of the stochastic controller was overly optimistic.

We were able to collect enough data for tests one and four, with only the former being a test measured by a quantitative metric. A thorough analysis of the stochastic controller decision behavior has not been completed, however we were able to analyze the data from test one, which was designed to validate the linear

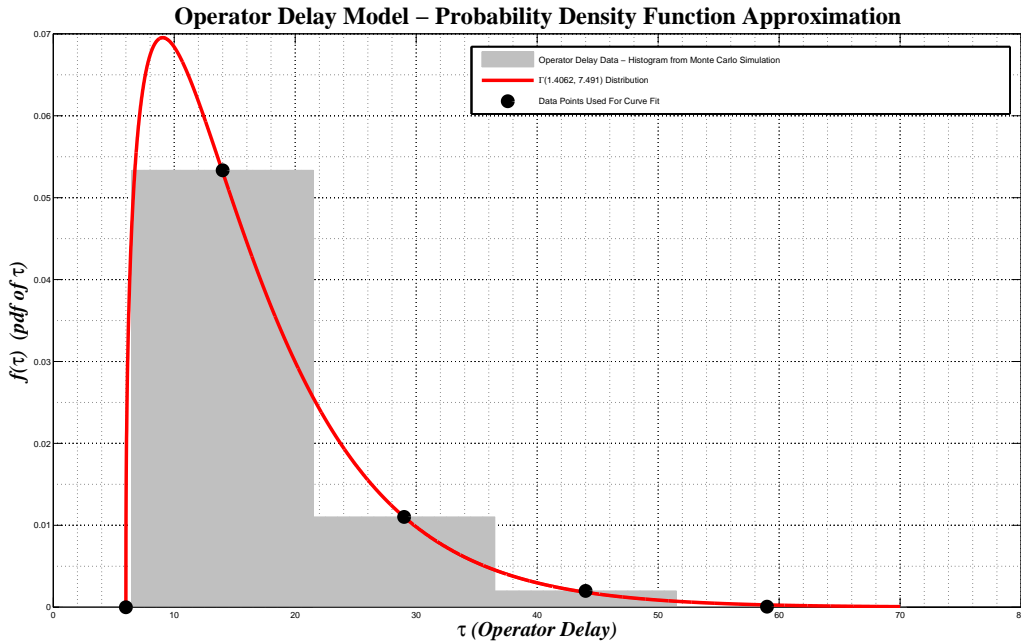


Figure 7. Operator Delay Model - Probability Density Function Approximation

model for the revisit cost. As previously mentioned, we collected the amount of reserve used for a range of operator delay values. The linear model of this cost is given by equation (27). The data were plotted and a least squares fit of the data was performed to fit a line to the data. This least squares fit is shown in Figure 8. The equation of the linear fit of the data is given by

$$\hat{J}(\tau_i) = 1.4\tau_i + 85. \quad (27)$$

We consider this to be a very promising result considering the limited amount of data that was able to be collected and keeping in mind that the theoretical model does not factor wind into the cost at all.

V. Conclusion

The effectiveness of the stochastic controller depends strongly on the reward function. To determine the effectiveness of a variety of reward methods, hundreds of simulation trials were performed and the resulting data were analyzed by a scoring algorithm. The results of the scoring algorithm seem to indicate that the quality of a particular reward function is related to the amount of saturation appearing in the threshold surfaces for that method. Although threshold surface saturation is not necessarily bad, it is not optimal. The method where saturation was mostly avoided was the method that performed the best. It was also shown that a discrete reward method can outperform an information theoretical method.

Essentially, the reward function outputs two values, expected rewards for the cases where the operator responds true or false. It may not be necessary to have a function dedicated to the determination of these values based on probabilities in the scenario. Instead, the two expected reward values could be determined using an optimization program that modulates the values over a series of simulations.

Although the flight test of the stochastic controller yielded little data, there were several important lessons learned from this particular flight test of the COUNTER program. The biggest lesson would be one of preparation. All of the tests should have been thoroughly executed multiple times successfully in simulation prior to the flight test. This was only partially accomplished due to scheduling constraints for the test range, thus time became an issue. Another lesson learned was that reliability of the platform on which the decision logic is being tested can seriously affect the success of the flight test, especially when it is pivotal that a MAV complete an entire mission for any of the data from that run to be useful.

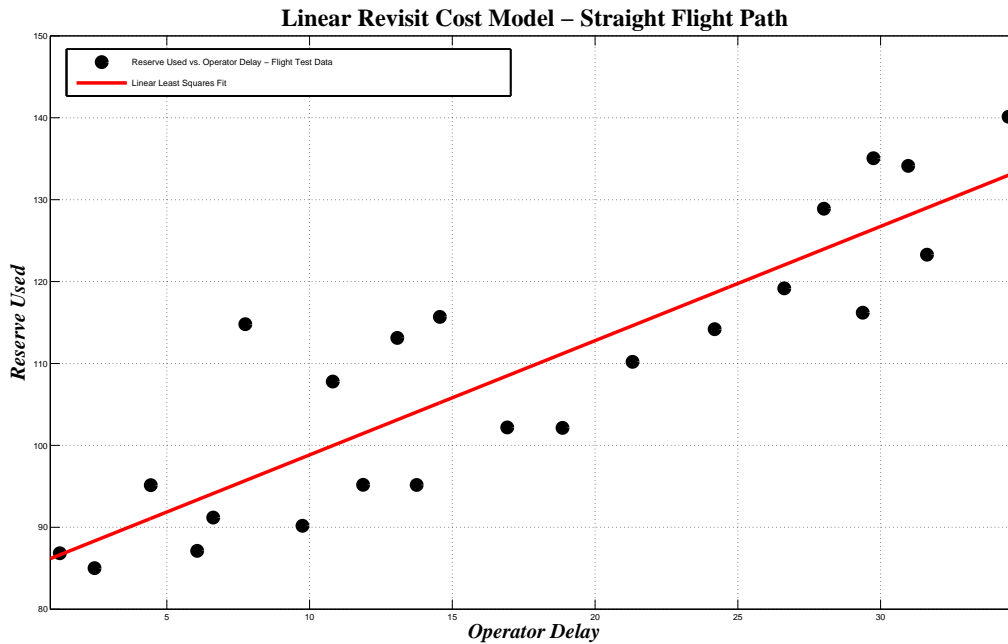


Figure 8. Linear Revisit Cost Model - Straight Flight Path

Plans are to attempt another flight test of the stochastic controller at a future flight test, time permitting. It is very likely that much more useful data would be collected for analysis.

References

- ¹Gross, D., Rasmussen, S., Chandler, P., and Feitshans, G., "Cooperative Operations in Urban Terrain (COUNTER)," *Proceedings of the SPIE*, Vol. 6249 of *Defense Transformation and Network-Centric Systems*, Orlando, FL, May 2006.
- ²Pachter, M., Chandler, P., and Darbha, S., *Cooperative Networks: Control and Optimization - Proceedings of the 6th International Conference on Cooperative Control and Optimization*, chap. Optimal Control of an ATR Module Equipped MAV-Human Operator Team, Edward Elgar Publishing, Northampton, MA, 2006.
- ³Pachter, M., Chandler, P., and Darbha, S., "Robust Control of Unmanned Aerial Vehicles in Uncertain Environments," *International Journal of Robust and Nonlinear Control*, Vol. 18, No. 2, January 2008.
- ⁴Pachter, M., Chandler, P., and Darbha, S., "Optimal Sequential Inspections," *Proceedings of the IEEE Conference on Decision and Control*, San Diego, CA, December 2006.
- ⁵Girard, A., Pachter, M., and Chandler, P., "Optimal Decision Rules and Human Operator Models for UAV Operations," *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Keystone, CO, August 2006.
- ⁶Girard, A., Darbha, S., Pachter, M., and Chandler, P., "Stochastic Dynamic Programming for Uncertainty Handling in UAV Operations," *Proceedings of the American Control Conference*, New York, NY, July 2007.