NUREG/CR-1140 Vol. I
UCRL-52712 Vol. I

AGGREGATED SYSTEMS MODEL OF NUCLEAR SAFEGUARDS

EXECUTIVE SUMMARY

POOR ORIGINAL

Prepared for
U. S. Nuclear Regulatory Commission
by
Lawrence Livermore Laboratory

PDR

8003200 346

Available from

GPO Sales Program
Division of Technical Information and Document Control
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

and

National Technical Information Service
Springfield, Virginia 22161

NUREG/CR-1140 Vol. I
UCRL-52712 Vol. I

AGGREGATED SYSTEMS MODEL OF NUCLEAR SAFEGUARDS

EXECUTIVE SUMMARY

Manuscript Submitted: April 1979
Date Published: February 1980

Prepared for
Research Division

U.S. Nuclear Regulatory Commission
Washington, D. C. 20555
Under Interagency Agreement DOE 40-550-75

NRC FIN No. A-0115

by
Lawrence Livermore Laboratory
Livermore, CA 94550

operated by University of California

for the U.S. Department of Energy

FOREWORD

This two-volume report describes the Aggregated Systems Model (ASM), a formal aid for nuclear safeguards decision making. This tool permits decision makers to integrate various forms of safeguards information (adversary characteristics, safeguard system effectiveness, costs of safeguarding, consequences of diverted special nuclear material) to provide an evaluation and ranking of complex safeguards alternatives.

The work reported here had its origins in studies for the Nuclear Regulatory Commission begun in 1977 at LLL by John Lathrop, Stein Weissenberger, and Ivan Sacks, with subcontractors Bruce Judd of Applied Decision Analysis, Inc. (ADA) and Rex Brown of Decisions and Designs Inc. During this period, the basic ideas and structure were worked out for a highly aggregated model of safeguards decision making, to provide a tool for organizing the analysis of this very complex problem.

In 1978, the concepts and models were further developed by Weissenberger at LLL and Bruce Judd and Jean Huntsman of ADA; the major thrust of this period was to refine the model and examine value-impact tradeoffs in evaluating and ranking decision alternatives. The most recent effort, which has added significant features and which forms the immediate substance of the work reported here, has been carried out by Rokaya Al-Ayat of LLL and Judd and Huntsmann of ADA.

This report consists of two volumes. The first volume--Executive Summary--summarizes the methodology and introduces some of the results that have been achieved. The second volume describes in detail the Aggregated Systems Model.

Stein Weissenberger

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

When setting the performance criteria for systems which safeguard special nuclear material (SNM), decision makers must consider characteristics of the adversaries who attempt to divert SNM, safeguards responses to these attempts, costs of safeguards systems, and the consequences of diverted SNM.

This report describes an Aggregated Systems Model, which is designed to assist decision makers in integrating and evaluating these diverse factors consistently. The report summarizes the results obtained from applying the model to safeguards decision making in areas such as the hardware or procedures required, substitution of electronics for human safeguards, and overall performance criteria for safeguards systems. New performance criterion designed to measure how safeguards systems deter adversary attempts are also described.

INTRODUCTION

The objective of the Nuclear Regulatory Commission's safeguards program is to protect the public from the risk of death, injury, or property damage due to malevolent uses of special nuclear material (SNM). Lawrence Livermore Laboratory (LLL) is developing analytical procedures to assist the NRC in meeting this objective.

This report describes an analytical tool developed by LLL and Applied Decisions Analysis, Inc. (ADA). This tool is designed to help the NRC set performance criteria for safeguards systems in facilities that handle SNM. The tool integrates diverse information about adversary threats, safeguards systems, and the consequences of diverting SNM. This helps decision makers to explore sensitivities and evaluate different assumptions in the criteria setting process. The tool can also help facility designers choose safeguards measures to meet the NRC's criteria.

In addition, the tool is useful for evaluating security at a given facility. It can provide a "first cut" assessment of the safeguards system's performance, the result of which can guide more detailed assessment. Moreover, if a detailed assessment* is performed, this tool can be used to incorporate the results into an overall measure of the safeguards system's effectiveness.

DECISION ENVIRONMENT

An analytical tool for safeguards decision making must be designed to integrate a wide range of factors. For instance, when evaluating the risk posed to a facility that handles SNM, the decision makers must consider many possible threats, the safeguards system's response to those threats, and possible consequences of a successful diversion of SNM. In the following

---

*LLL is developing assessment procedures to assist the NRC in determining Licensee compliance with safeguards regulations.

paragraphs we describe briefly four of the factors that are inherent in the safeguards decision problem and indicate the advantages and limitations of using the analytical approach to deal with them.

Uncertainty

The adversary's motives, resources, knowledge, and plan of attack are all uncertain. Because safeguards systems are not perfectly reliable and are not immune from tampering, their ability to thwart an attempt is also uncertain. In addition, the consequences of the use of diverted SNM cannot be estimated precisely.

Data

The regulator must rely on a mix of objective technical data and subjective judgment when setting criteria. Safeguards system performance characteristics can be determined to some extent, but adversary characteristics and the consequences of the use of diverted SNM are still in the domain of expert opinion.

Balancing Risk and Cost

The regulator must consider the trade-off between the cost of reducing the risk and the social benefits of risk reduction when setting criteria.

Communications

Information must flow from scientists and engineers to regulators, to law makers, and to the public. Safeguarding SNM is a technological problem that involves the use of highly specialized experts in fields ranging from law enforcement to physics. Safeguarding SNM is also a social problem, which should incorporate the public's perceptions and values.

ADVANTAGES AND LIMITATIONS OF THE ANALYTICAL TOOL

The analytical tool described here helps the decision maker who confronts the following factors:

- Uncertainty--The tool considers uncertainty explicitly. Using a dynamic and probabilistic model, the tool helps the decision maker investigate the system's vulnerability to many types of adversaries, some of whom may make repeated covert attempts to divert SNM.
- Data--We will never have precise data on all adversaries and consequences, or on performance. Because judgmental input is required, the tool does not yield exact answers to these safeguards questions. However, the tool does make use of existing performance data, whenever it is available. The tool also shows which judgmental data have the greatest influence on the safeguards decisions and where additional research is needed.
- Balancing risk and cost--The tool improves the consistency of safeguards criteria by quantifying value judgments on both sides of the benefit-cost trade-off, thus enabling decision makers to strike a balance. Through sensitivity analysis, the tool identifies the value judgments that are most crucial to the criteria setting decision.
- Communications--The tool is designed to assist communication among parties during the decision process. Although the tool requires some technical inputs, it is sufficiently aggregated and simple that nontechnical people can understand the link between technical inputs and model outputs. Therefore, the tool can play an important role in the process of setting and justifying regulations. The tool also documents major assumptions and technical judgments behind the decision and allows them to be checked. As information changes, the tool can be used to determine any appropriate changes in safeguards criteria.

The tool is not a computer formula for making decisions; it aids the human decision makers by reflecting their judgment, not replacing it. While some people might question the wisdom of making explicit, social value judgments in a political environment, we believe that the consistent application of value judgments, the insight gained through explicit analysis, and the logical justification of regulation can greatly enhance the decision-making process. In Ref. 1, Ralph Keeney discusses further the merits of explicit treatment of value judgments.

EXAMPLE OF SPECIFIC SAFEGUARDS DECISIONS

The analytical tool is useful for making safeguards decisions at several different levels. The tool can help set performance requirements for individual components, for subsystems such as material control or physical security, or for the entire safeguards system at a facility. In addition, the tool can be used to help a designer choose which components to include or exclude, or to evaluate the benefits when substituting one component for another.

We will illustrate the spectrum of uses for setting performance criteria for a nuclear facility and show the results obtained from analyzing three sample decisions. This type of analysis for setting criteria is often called value-impact analysis. We will also show the results of a value-impact analysis for a number of facility designs. The three sample decisions are as follows:

- Require a two-person rule for all plant operations.
- Substitute a real-time SNM quantity estimator for frequent physical inventory.
- Set a constraint on overall facility safeguards performance in terms of a limit on the expected quantity of SNM diverted per operating year.

As mentioned above, the tool can be used for assessing the facility's compliance to a given performance criteria and for aggregating the results of the detailed assessment of a facility. (See Refs. 2, 3 and 4.) Although we do not demonstrate these capabilities in this summary, these uses are described in Ref. 5.

AGGREGATED SYSTEMS MODEL

OVERVIEW

The tool we have developed to analyze the previously mentioned safeguards decisions is called the Aggregated Systems Model (ASM). We use the term "aggregated" because the model is a combination of highly aggregated models, and the term "system" because it covers a broad perspective, representing many of the major factors in setting safeguards criteria.

The ASM is a collection of submodels as shown in Fig. 1. The submodels include the following:

- Adversary Model--A probabilistic model of several kinds of adversaries, which incorporates attempt frequency, resources, and plan of attack (called diversion strategy), including how many tries it will take before they have stolen the desired quantity of SNM.
- Facility Model--This model describes the likelihood of each adversary being detected and interrupted by the safeguards system.
- Safeguards Technology Model--This model describes safeguards components that could be incorporated in the safeguards design. Ultimately, data for both the Facility Model and Safeguards Technology Model will be supplied by detailed analyses of facilities.
- Consequence Model--This model is a probabilistic statement of the effects of malevolent uses from diverted SNM.
- Utility Model--Outcomes such as diversion consequences and safeguards costs are expressed in common units using this model. The model allows the decision maker to make explicit trade-offs among lives lost, property damage, and evacuation costs, which collectively are called diversion costs (C_D), and the cost of safeguards (C_S). Their sum ($C_S + C_D$) is the total cost (C_T).

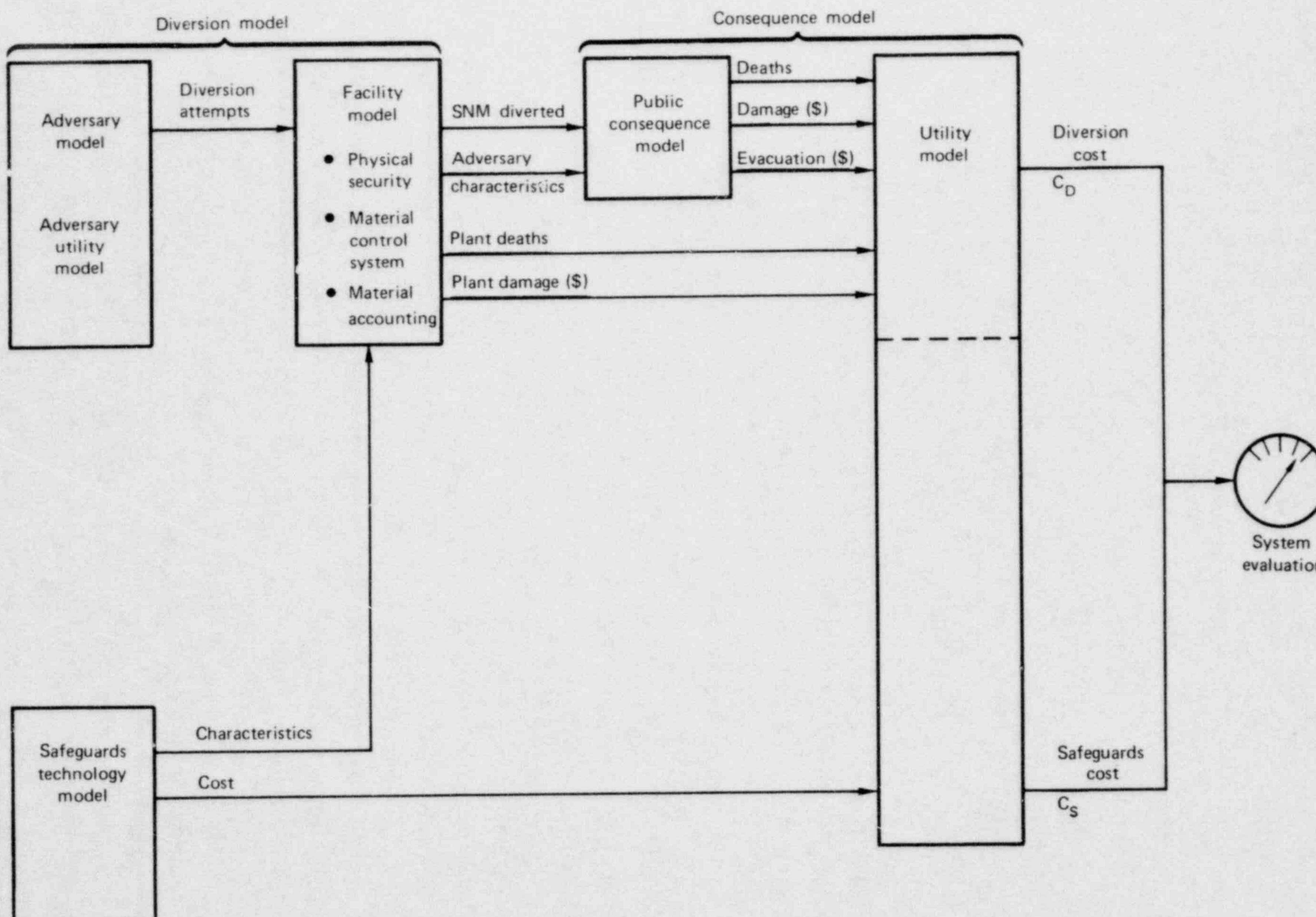


FIG. 1. Aggregated systems model overview.

CAPABILITIES AND USES OF THE ASM

We can use Fig. 1 to conceptualize how the ASM will help analyze the types of safeguards decisions mentioned earlier. System design or operating procedure decisions such as the two-person rule or substituting real-time quantity estimators for inventories require at least the Facility and Safeguards Technology Models. These models will produce measures of facility performance against a specified adversary threat and show how performance changes as the design or operating procedures are varied. When the Adversary Model is coupled with the Facility and Technology Models, one can see how the facility performs against the full spectrum of adversaries, all of whom have optimized their plan of attack against the facility. The same three models are used to determine the overall facility performance, as measured by the expected quantity of SNM diverted each year.

Although the models discussed so far can be used to measure the safeguards systems performance, they are not adequate when determining the optimum level of performance. Although the models can assess the performance against a criterion, they cannot determine the best level at which to set the criterion. Setting the criterion requires a comparison of the cost (impact) with the reduction in risk (value) for different safeguards criteria. The analytical procedure is often called value-impact analysis.

The first step of the value-impact analysis is to analyze how the designs or requirements affect the facility's performance against adversaries. In Fig. 1, this analysis is accomplished in the box labeled "Facility Model." Performance is measured by the arrow labeled "SNM Diverted," which emanates from the Facility Model. If the decision maker can intuitively balance cost (impact) with the benefits of reduced risk of SNM diversion (value), then the value-impact analysis stops here. However, if the decision maker feels that additional quantitative analysis would be helpful, then the analytical process continues with step two.

The second step is used to express the public consequence of diverting the SNM. For example, the ASM can describe how tightening security reduces the amount of SNM diverted and therefore reduces the potential number of deaths or damage due to malevolent use of SNM. This translation from quantity diverted to consequences occurs in the Consequence Model box shown in Fig. 1. Once again, the process can stop, and at this point the decision maker implicitly balances the cost of safeguards with reduced consequences. However, step three can aid this balancing process.

The third step involves balancing changes in public consequences (values) with the social costs (impacts) of the safeguards designed to reduce consequences. The Utility Model uses explicit value judgments, which are applied consistently across all consequences. The model also shows the sensitivity of various conclusions to changes in important value judgments.

As an aid to detailed decision making, the ASM can use some, or all, of the models shown in Fig. 1. When determining whether a facility meets design criteria, the decision maker needs at least the Safeguards Technology and Facility Models, especially when evaluating performance against a specified adversary threat. The Adversary Model is used to test facility performance against an entire range of threats. The output of such an analysis is a statement of the given facility's performance measured in terms of probabilities of detection, interruption, or the quantity of SNM diverted.

When the ASM is used to set performance criteria, more of the models in Fig. 1 may be used. If the goal is to show which facility components are the weakest links in safeguards, only the Facility and Adversary Models are needed. Assessing the total social risk posed by the plant requires the addition of the consequence model. Choosing the socially optimal level for performance criteria involves tradeoffs which are aided by the Utility model. Thus, at least conceptually, the decision maker would consider all the factors shown in Fig. 1 when setting criteria.

SAMPLE RESULTS

Before describing the models and their interaction, we will illustrate two typical results of the analyses. The data for the example analyses are purely illustrative, and were developed by a few members of the LLL Safeguards Project team. These data should not be regarded as accurate, as they are provided only as an illustration.

First, we examine the decision to set a constraint on the expected quantity of SNM diverted each year. As mentioned above, setting a criterion will utilize all three analytical steps outlined. The decision criterion is demonstrated graphically in Fig. 2. The horizontal axis measures system performance for 14 different system designs. We measure overall performance in terms of expected quantity of SNM diverted per year. Assume that the regulator is setting a standard for this quantity. On the vertical axis, we measure three costs as a function of expected SNM for the 14 designs: safeguards cost (C_S), diversion cost (C_D), and total cost (C_T). The optimal level of performance is where total social cost C_T is lowest. This level occurs when 6 g/y are diverted. To the left of 6 g/y, C_D is lower, but the cost, C_S, of this decrease is substantial, causing C_T to rise rapidly.

Figure 3 illustrates the results of a value-impact analysis for the 14 designs chosen. The figure is a simplified representation of the relationship between diversion cost, C_D (value), and safeguards cost, C_S (impact). Each point on the graph corresponds to a different system design. Those that are on the boundary curve are most efficient for achieving given performance at minimum cost. The curve illustrates that the risk of diversion can be lowered, but at increasingly high safeguards cost, C_S. The curve shows an NRC commissioner or member of Congress, for example, how much it costs to reduce diversion risks to a particular level. Note that for the illustrative data set, it becomes very expensive to reduce diversion cost below about $\$3 \cdot 10^6$ per year.

The choice of an optimal safeguards design depends on the decision maker's trade-off between safeguards cost (C_S) and diversion cost (C_D). The slope of the straight line in Fig. 3 is minus 1, which indicates a 1 to 1 relative weighting between C_S and C_D. The point of contact between the curve and the efficient boundary line gives the optimal system--Design D6 in this case. For the designs depicted, Design D6 will be optimal over a wide range of weights. If the weight assigned to C_D by the decision maker is anything more than 38 percent of the weight assigned to C_S, then Design D6 will be optimal. Notice that an extremely large weighting on C_D would be required before Design DM (maximal design) would be optimal.

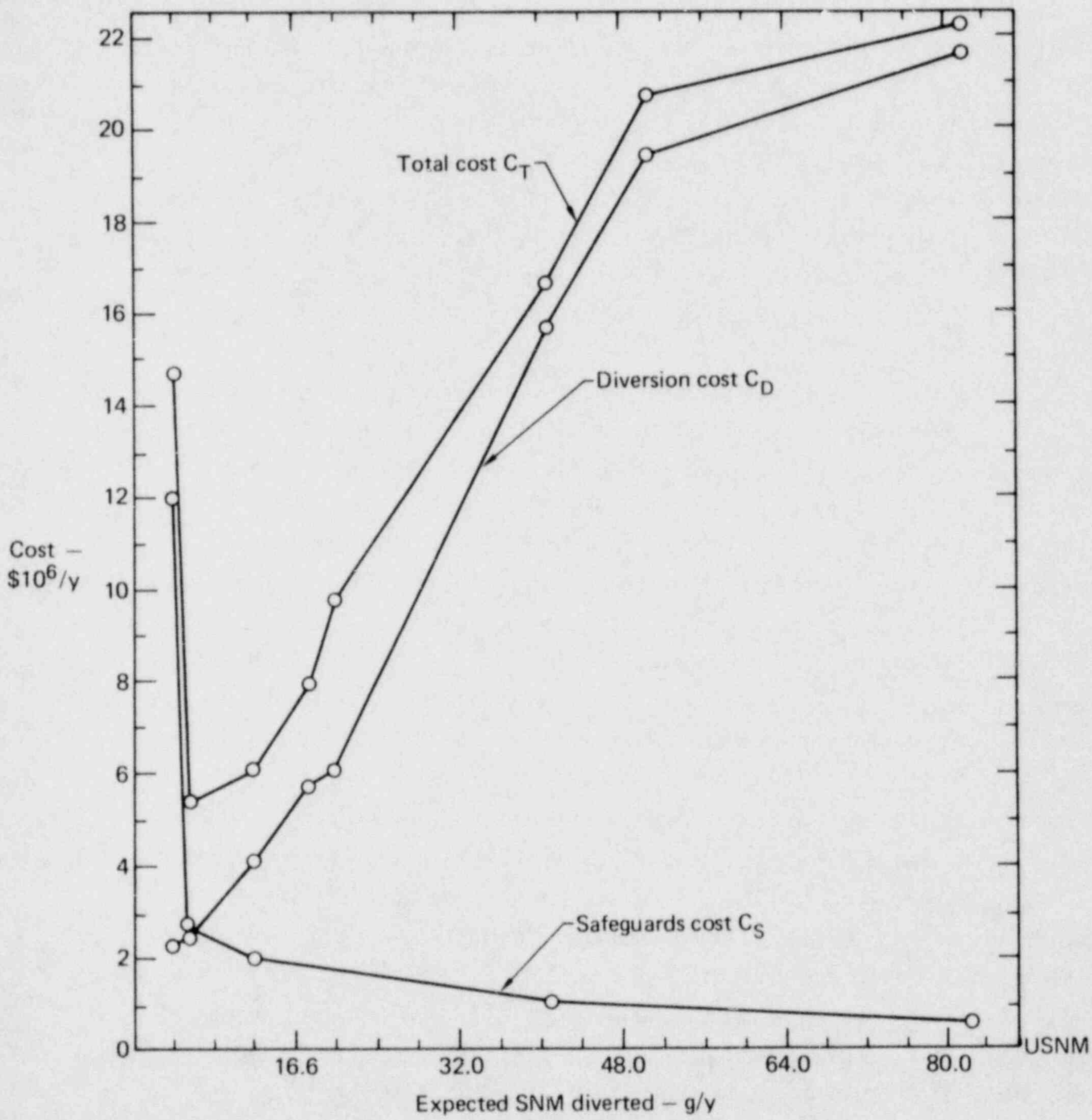


FIG. 2. Performance graph (cost vs expected amount of SNM diverted per y).

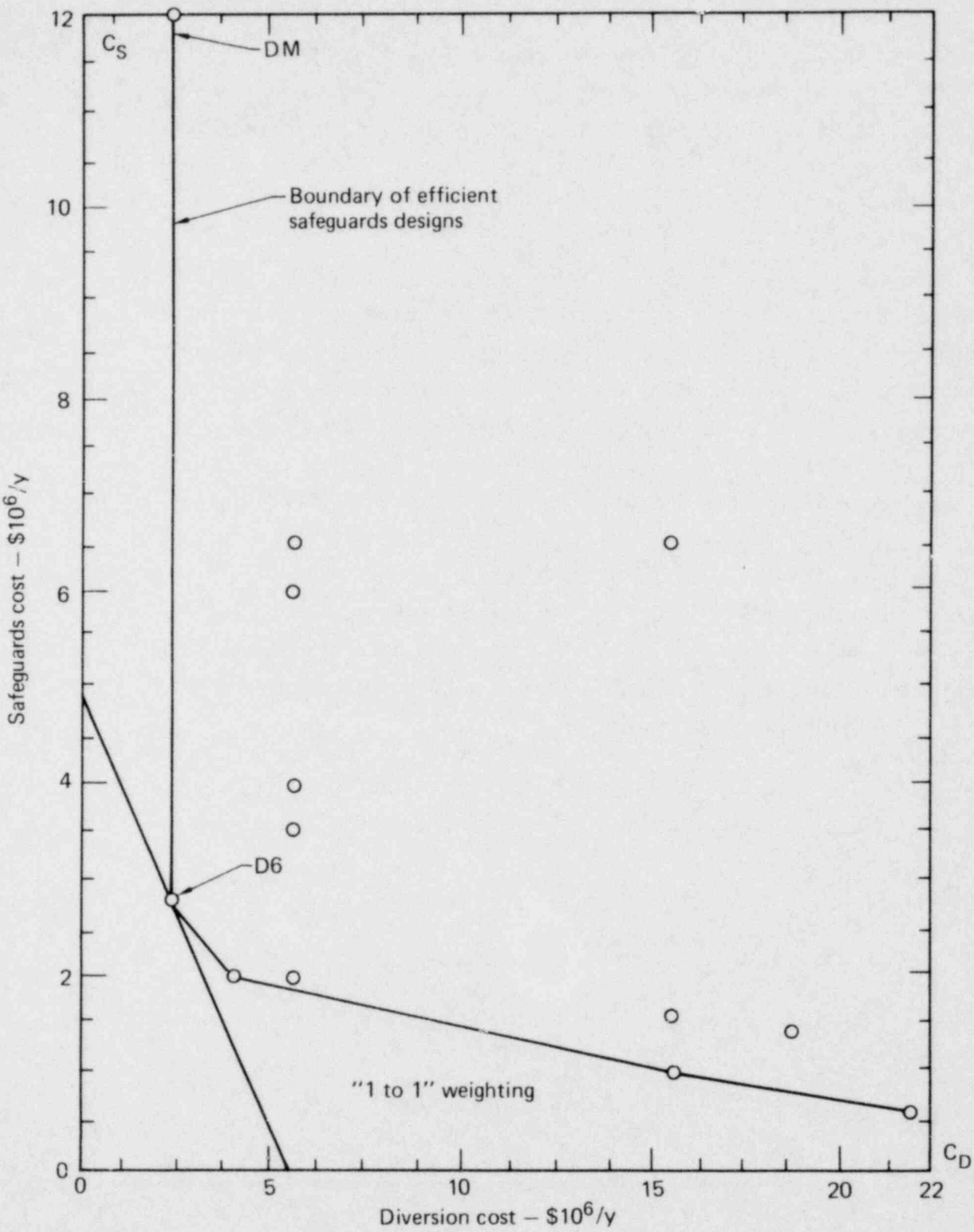


FIG. 3. Value-impact trade-off curve (safeguards costs vs diversion costs).

Similarly, this graph shows how the optimal design would change if numbers are varied in the Consequence Model. If the weighting between C_S and C_D is kept at 1 to 1, then C_D could be reduced by 38 percent or increased almost limitlessly, and Design D6 still would be optimal. Of course, the robustness of Design D6 most likely would be less if there were more than 14 designs plotted in Fig. 3.

The implication of this result is that the design decision is virtually insensitive to an increase in consequences. A doubling of the number of deaths and injuries resulting from a nuclear detonation in the Consequence Model would result in an 81 percent increase in the expected consequences of a 5-kg diversion; C_D for all designs increases by slightly less than 80 percent. This increase still does not make Design DM nearly as attractive as Design D6.

The design decision is similarly insensitive to changes in the trade-off weights assigned to C_S and C_D. In an actual analysis using real data, this could be an important insight for decision makers who must balance social risks with economics. The paper by Keeney (Ref. 2) further explores this benefit of explicit value analysis.

DESCRIPTION AND ILLUSTRATIVE APPLICATION OF THE AGGREGATED SYSTEMS MODEL

This section describes the models and interactions illustrated in Fig. 1. These are: (1) Facility Model, (2) Adversary Model, (3) Interaction of Facility and Adversary Models, (4) Consequence Model, and (5) Utility Model. The descriptions are followed by illustrative model results and sample analyses of the three previously mentioned safeguards decisions. The purpose of this section is to provide a general understanding of the kind of input data required and the types of output available. The data for the example analyses are purely illustrative and should not be regarded as accurate.

FACILITY MODEL

The Facility Model represents the system designed to safeguard SNM. The current version of the ASM includes nine basic building blocks (components) of this system. These components are listed in the lefthand column of Table 1.

When referring to the safeguards system design, we mean the list of components installed in the facility. Changing the design means adding, deleting, or substituting components. Six sample designs are shown in Table 1.

The performance of individual components determines the performance of the safeguards system. Calculating system performance requires: (1) input data on component performance, and (2) an algorithm for aggregating component performance into safeguards system performance. One step in the algorithm is to aggregate component performance to an intermediate level called "subsystem" performance. Subsystems are collections of components that have similar detection characteristics. For instance, in the current version of the ASM, we aggregate components into three subsystems as follows:

- electronic detection
- visual detection
- accounting (records detection).

TABLE 1. Components of safeguards system designs.

Components	Sample Safeguards Designs D1 Base case	D2 Minimal safeguards	D3 Two-person rule	D4 Quantity estimators	D5 Frequent inventory	D6 Full physical security	Subsystems
1.1 Quantity estimators				X			Electronic Detection
1.2 Process state monitors	X		X	X	X	X	
1.3 Personnel monitors	X		X	X	X	X	
1.4 Procedure monitors							
2.1 Stationary guards	X	X	X	X	X	X	Visual Detection
2.2 Roving guards						X	
2.3 Two-person rule			X			X	
3.1 Nominal accounting	X	X	X	X	X	X	Accounting
3.2 Frequent inventory					X		

Accounting components include a required basic accounting system and a frequent (bimonthly) inventory of SNM at the facility.

The aggregation scheme is shown on the right side of Table 1.

ADVERSARY MODEL

This subsection describes the following:

- The types of adversaries currently in the model.
- The decisions adversaries make about how to divert SNM.
- The values adversaries assign to the possible outcomes of the attempt.

Table 2 lists the characteristics of the 14 adversary types we initially considered. Sensitivity analysis of the illustrative data set shows that six of these adversaries dominated the other eight in terms of their contribution to the expected quantity of SNM diverted per year (percentage contributions are shown on the right in Table 2). Therefore, only those adversaries that are shown by asterisks in Table 2 were retained for this analysis.

We define an attempt for all types of adversaries as the existence of an adversary who wants to divert SNM. The attempt may consist of one or multiple tries, or, if the adversary is deterred, no tries at all.

It is assumed that the various adversary types will optimize their plan of attack depending on the type of safeguards in the facility. We call this adversary choice a strategic decision. To make a strategic decision, the adversary considers the following:

- Whether or not it is worth the risk to try to divert SNM.
- The quantity of SNM stolen each try (increment size).
- The number of tries.
- The diversion path (route to and from the SNM), which is defined in terms of which components could detect the diversion attempt.

We define adversary tactical decisions as those involving a decision to abort a try if detected by a safeguards system. Presumably, aborting might

TABLE 2. Adversary characteristics.

		Resources						
Number	Access	Equipment	Authority	Collusion	Desired quantity	Attempts	Attempt frequency (Att./1000 y)	Percent of SNM diverted
1*[a]	Outsider	Major	No	No	Bomb	One	2	38
2	"	Minor	"	"	"	"	0.02	0.003
3*	Insider	Major	"	Yes	"	"	0.29	4.7
4*	"	"	"	"	"	Multiple	2.6	33
5	"	"	"	"	Less	One	0.03	0.049
6	"	"	"	No	Bomb	"	0.03	1.1
7	"	"	"	"	"	Multiple	0.26	1.2
8	"	"	"	"	Less	One	0.03	0.006
9*	"	Minor	"	Yes	"	Multiple	13	13
10	"	"	"	No	"	"	1.4	0.14
11*	"	Major	Yes	Yes	Bomb	One	0.03	1.2
12*	"	"	"	"	"	Multiple	0.15	7
13	"	"	"	"	Less	One	0.07	0.11
14	"	"	"	"	"	Multiple	0.11	0.18

[a]Asterisks indicate dominant adversary types.

reduce the chances of being identified, as the safeguards system might assume the detection is a false alarm.

We assume that adversaries make strategic and tactical decisions based on their own preferences for the ultimate outcome of the attempt. These outcomes are as follows:

- complete success
- partial success
- failure (no success or status quo)
- capture.

Preferences for these uncertain outcomes are represented in a utility function. For example, we hypothesize the following utility function for a highly sophisticated insider with major equipment, who is in collusion with other insiders.

Utility Function

Outcome	Utility
Complete success	1.0
Partial success	0.01S + 0.76
No success (or no try)	0.[illegible]
Capture	0.[illegible]

Where S is the quantity of SNM diverted

$$0.3 \text{ kg} \leq S \leq 10.0 \text{ kg}.$$

We assume the adversary evaluates all possible strategies and tactics by considering the likelihood and utility of all outcomes. The adversary then chooses the strategy or tactic with the highest expected utility. We also assume, for this example, that the adversary knows exactly how the safeguards system will respond to an attempt.

INTERACTION: FACILITY AND ADVERSARY MODELS

We have modeled the complex interaction between the adversary and the facility safeguards in some detail. Two factors make the interaction complex. First, the adversary will optimize his or her strategy to find the weakness in the safeguards system. As new safeguards are installed to close off one diversion path, we assume the adversary will choose the next best. The next best option may be to make no try to divert the SNM, in which case the adversary is deterred.

Assuming the adversary is not deterred, he or she may make repeated covert tries. A safeguards component that is "slow but sure" to detect, such as an accounting system, may be valuable in preventing diversion by this type of adversary.

Figure 4 is a schematic decision tree that shows major decisions and outcomes in the adversary-facility interaction. Square nodes represent decisions. (An "S" in the box denotes decisions made by the facility designer following NRC guidelines in the best interest of society, and an "A" denotes adversary decisions.) Circles represent probability nodes, where the decision maker is uncertain about the outcomes. We show only two branches at each node, although there can be many.

The first decision we consider is the facility design (which safeguards components to choose). Obviously, this decision is made without perfect foresight about which adversaries (if any) will attack and whether the system can detect or interrupt the attempt. The three sample decisions listed previously are special cases of this general design decision.

The second node in Fig. 4 represents the spectrum of adversary threats. The choices of strategy and tactics made by these adversaries are represented by the third node. These decisions consider the probabilities of detection, identification, and interruption, which lead to the outcome states shown on the righthand side of the figure.

Facility safeguards response to the threat is measured in terms of the probabilities of detection, identification, and interruption. Detection means

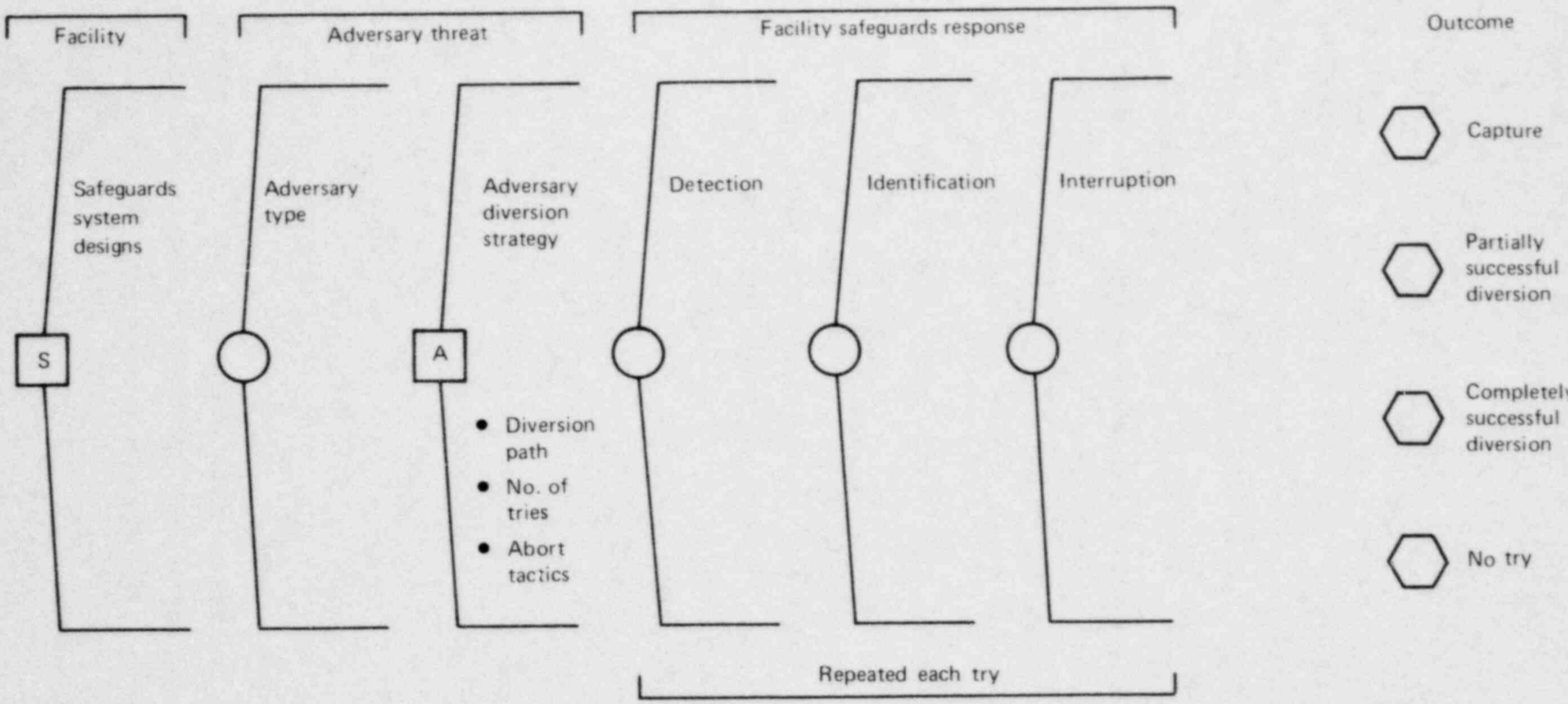


FIG. 4. Decision tree (adversary-facility interaction).

that the safeguards authority receives an alarm, but the signal is not distinguishable from a false alarm. Identification means the authority knows both the adversary's identity and that the alarm is real. In this case, the guard force can be dispatched to confront the adversary. Interruption means that no SNM crosses the plant boundary.

Some adversaries face the detection/identification/interruption sequence of events several times on repeated trials. Also, we assume that the adversaries know the probabilities of detection, identification, and interruption by the safeguards system.

The ASM translates input data on component and system detection, identification, and interruption probabilities, as well as information on adversary strategy and tactics, into a probability distribution for the attempt outcomes shown in the schematic decision tree.

Figure 5 shows the results of this calculation for one adversary. Here, the decision is the adversary's strategic choice of SNM increment size. The system design, shown as D2 in our illustrative data set, has minimal safeguards. The adversary is a highly sophisticated insider with good equipment, who is in collusion with other insiders. The desired quantity of SNM is 10 kg. It is assumed that this adversary can try repeatedly, and will consider three increment sizes: 10 kg; 1 kg; or 300 g. The fourth option is to make no attempt to steal SNM.

The figure shows that the most likely outcomes in cases where the adversary makes the try are either capture or complete success. These outcomes are valued according to: (1) how much SNM is diverted; (2) the utility functions described earlier. The strategy with the highest expected utility (0.89 = (0.894) (1.0) + (.001) (.7)) in Fig. 5, also has the highest probability of complete success and the greatest expected quantity diverted. We assume the adversary would choose "one large quantity" for this particular design.

In contrast to the design used for the strategy choice in Fig. 5, all other designs considered in our illustrative data set had more safeguards, and for half of the designs the adversary was better off not to make an attempt; in other words, for these designs the "No Try" option had the highest expected utility.

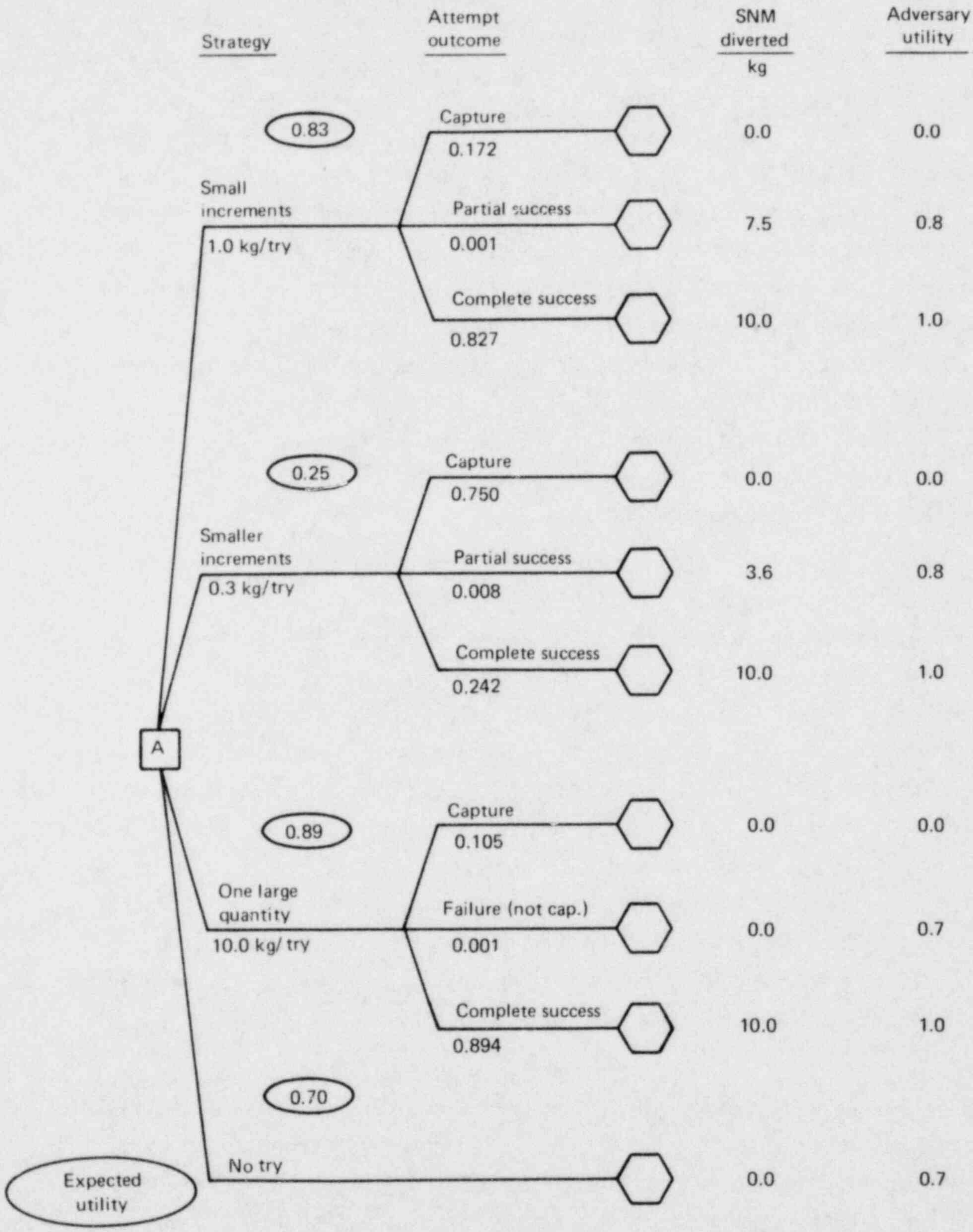


FIG. 5. Sample strategic choice (adversary 4--design D2).

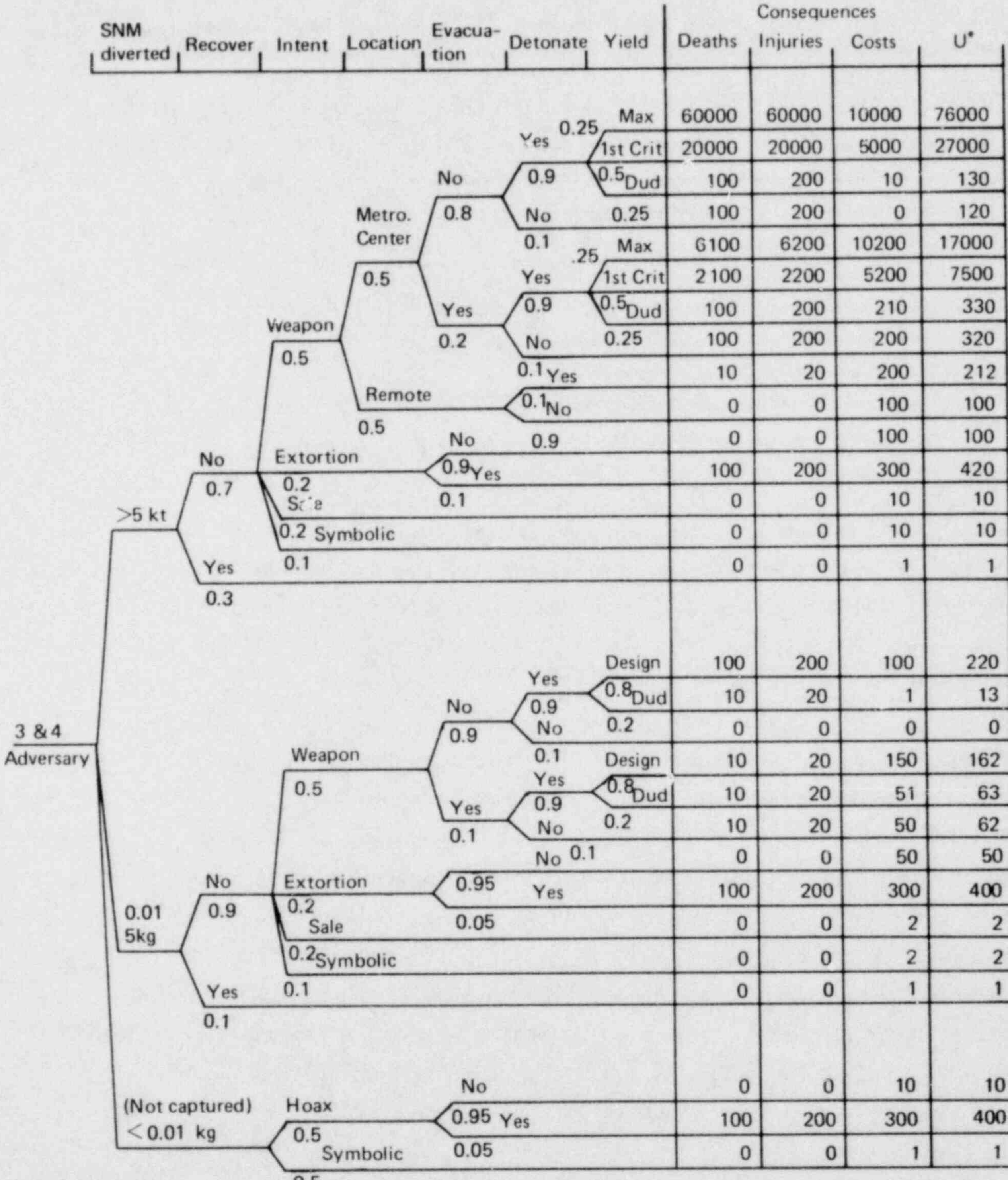
ILLUSTRATIVE PUBLIC CONSEQUENCE MODEL

Deciding whether the social benefits of a safeguards design are worth the social costs requires a judgment regarding the social consequences of the diversion of SNM. Usually this judgment is performed implicitly. However, explicit consequence modeling can improve consistency in this decision process. Possible sequences of consequence model events may be evaluated in the probability tree framework shown in Fig. 6. Uncertain events in the Public Consequence Model are as follows:

- recovery of the material by law enforcement agencies
- adversary's intended use of SNM, such as building a weapon, extortion or sale
- location where any weapon is to be detonated
- whether or not a threat was received and evacuation occurred
- actual detonation of weapons
- yield from any weapon.

The decision trees in Figs. 4 and 5 produce probability distributions on SNM diverted. The probabilties from these distributions are placed on the first branches in Fig. 6. A probability tree like Fig. 6 is made for each adversary, and probabilities of the above events are assigned. For example, if adversary 4 diverted 10 kg of SNM, the model shows a 30 percent chance that it will be recovered. There is a 50 percent chance that the adversary will try to make a weapon of it, and a 50 percent chance that the adversary will attempt to detonate the weapon in a metropolitan area. The illustrative data then shows that there is an 80 percent chance that there will be no evacuation, a 90 percent chance that the device will detonate, and a 50 percent chance that it will reach maximum yield.

The illustrative consequences of the events are measured in deaths, injuries, evacuation cost, and property damage. Weapons consequences are purely conjectural. However, a report soon to be issued by Dr. Dean Kaul at Science Applications, Inc. will describe a model which will predict weapons consequences.



*Utility Function:

1 × Deaths + 0.1 × Injuries + 1 × Costs = Utility (Equiv. $\$10^6$)

FIG. 6. Illustrative consequence model.

UTILITY MODEL

Measuring the consequences in common units is helpful for making implicit or explicit trade offs between safeguards cost and diversion risk. In the Public Consequence Model, we evaluate outcomes in terms of public fatalities, damage, dollar loss, and evacuation costs. We have selected dollars as the common unit of measurement. Illustrative trade-off values in dollars are \$100,000 per injury and \$1,000,000 per death. In other words, the judgment implies society is willing to pay $\$10^6$ to prevent the loss of one statistical life and $\$10^5$ to prevent one statistical injury. Evaluation of each outcome in these utility units is given in the righthand column of Fig. 6.

SAMPLE RESULTS

Before summarizing the results that pertain to the three sample decisions, the results of the adversary's decision process should be highlighted. Table 3 shows the results for the base case design and the minimal safeguards design. (See Table 1.) Notice that all adversaries will attack Design D3, but most are deterred by Design D1. This is because the adversary's chances for success are much lower with Design D1, so the adversary chooses not to risk being caught in an attempt.

Next, we consider a sample decision: whether or not to require a two-person rule for all plant operations involving material movement. The ASM will produce several measures of performance by which two-person rule may be judged. These measures are listed in Table 4 along with performance comparisons of designs with and without the two-person rule. The model shows, as expected, that requiring the two-person rule increases both the probability of detection and probability of identification, but the probability of interruption decreases. The decrease is due to the fact that most adversaries have been deterred, and only the outsider adversary--adversary 1, against whom the system has low probability of interruption--makes an attempt to steal SNM.

We assume that the cost of the two-person rule is \$500,000 per year. This cost includes 10 extra operators at \$50,000 total support cost. Compared to

TABLE 3. Comparison of adversary strategy and tactics for two designs.

	Adversary	Design D1		Design D3	
		Strategy	Tactics	Strategy	Tactics
1	Outsider	Assault	Never abort	Assault	Never abort
3	Insider - one attempt	No attempt	-	Working hours diversion	Never abort
4	Insider - multiple attempts	No attempt	-	Working hours 10 kg	Abort if guards detect
9	Insider - minor resources	No attempt	-	Divert from sampler	Never abort
11	Manager - one attempt	Disable monitors	Abort if guards detect	Working hours	Never abort
12	Manager - multiple attempts	No attempt	-	Working hours 10 kg	Never abort

TABLE 4. Evaluating the two-person rul

Performance measures	Design D1 Base case	Design D3 Required two-person rule
Subsystem performance		
Probability of detection	0.2	0.5
Probability of identification	0.2	0.4
Probability of interruption	0.2	0.1
System performance		
Frequency of attempt per 1000 y	17	3
Expected number of attempts	6	1
Expected SNM diverted (g/y)	42	18
Aggregated performance measures		
Diversion cost ($\$10^6$/y)	15.4	5.8
Safeguards cost ($\$10^6$/y)	1.7	2.2
Total social cost ($\$10^6$/y)	17.1	8.0

reduced consequences (valued at $15.4 - 5.8 = $9.6·$10^6$/y), the benefits of the two-person rule far exceed its costs. Although the data are purely hypothetical, this conclusion holds for an order-of-magnitude range of the trade-off values between lives saved and dollar cost. This conclusion is highly dependent on the data set and with real data the outcome might be different. Table 5 shows the results of an integrated safeguards evaluation in which a real-time quantity estimation system is substituted for frequent physical inventories. On the aggregated measures, that is, cost of diversion (C_D), cost of safeguards (C_S), and total cost (C_T), the design with quantity estimators is superior. The same holds true for system performance measures: frequency of attempt and the expected amount of SNM to be diverted. The design with quantity estimators (Design D4) has lower probabilities of detection and identification than the design with frequent inventories (Design D5). This is because adversary 4, the sophisticated adversary, is deterred by Design 4, but not deterred by Design D5. When adversary 4 tries against Design D5 he or she has a good chance (0.53) of being detected, which raises the overall probability of detection for Design D5. Notice, however, that a high probability of detection does not necessarily lead to the best overall system performance. (This is the same phenomena as shown by Table 4.)

The third sample decision, which involves setting a constraint on the expected quantity of SNM diverted each year, was discussed earlier and is illustrated by Fig. 3.

TABLE 5. Comparison of frequent inventory and quantity estimation monitors.

Performance measures	Design D5 Base case plus frequent inventory	Design D4 Base case plus quantity estimators
Subsystem performance		
Probability of detection	0.2	0.1
Probability of identification	0.3	0.2
Probability of interruption	0.2	0.2
System performance		
Frequency of attempt per 1000 Y	17	14
Expected number of attempts	6	7
Expected SNM diverted (g/y)	42	21
Aggregated performance measures		
Diversion cost ($\$10^6$/y)	15.4	5.9
Safeguards cost ($\$10^6$/y)	6.7	4.2
Total social cost ($\$10^6$/y)	22.1	10.1

SUMMARY

The Aggregated Systems Model (ASM) is a tool for setting performance criteria at a relatively high level. The model integrates many factors in these criteria decisions, including adversary information, facility safeguards response, consequence of using the SNM diverted, and value-impact trade-offs. The model is meant to be an aid to, not replacement for, the judgments of the regulators who must integrate these factors and make the ultimate decisions concerning safeguards.

Because of the uncertainties and complexities associated with the risk/cost trade-offs, quantitative models of the various factors involved in these decisions are necessary. The safeguards evaluation framework introduced here shows how various factors are interrelated and how the quantitative information about them is used to make a decision.

The analytic process has generated a set of performance measures to judge system effectiveness in thwarting various adversary types. The choice among these measures is determined by the type of system being evaluated. For example, when evaluating a Material Control and Accounting system, one needs to consider the probability of detection or the amount of SNM expected to be diverted; physical security might be evaluated using the probability of identification or capture.

A set of measures for the performance of the overall system is also generated by the analytical process. Of interest is the measure of deterrence, because it is rooted in the adversary's choice of how to attack a particular facility. We hypothesize that the adversary will not make the attempt if safeguards are strong enough. A quantitative framework for determining this measure of deterrence is presented in this paper.

Although an illustrative data set is presented, much of the information is judgmental. The ASM gives the regulator a tool to measure the sensitivity of the decision to this judgmental information. This insight can guide data collection efforts for the safeguards decision making process.

REFERENCES

1. Keeney, R. L., "Formal Inclusion of Value Judgments in the NRC Decision Making Process," Woodward-Clyde Consultants, June, 1979.
2. M. Dittmore, F. Gilman, J. Lim, Facility X Vulnerability Analysis: Phase I Report, NRC Report (classified) forthcoming.
3. A. Farziale and I. Sacks, "The Structured Assessment Analysis of Facility X," Volumes I and II (Volume II classified), Lawrence Livermore Laboratory, Livermore, Calif., Internal, January 1979.
4. T. Rice, and S. Derby, "Overview of the Structured Assessment Approach and Documentation of Algorithms to Compute the Probability of Adversary Detection," Applied Decision Analysis, Inc., December 1978.
5. "An Aggregated Systems Model for Nuclear Safeguards Decisions," Applied Decision Analysis, Inc., March 1979.

CB/nll

NRC FORM 335 (7-77)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER *(Assigned by DDC)*
NUREG CR/1140 Vol. 1
UCRL 52712 Vol. 1

4. TITLE AND SUBTITLE *(Add Volume No., if appropriate)*
Aggregated Systems Model of Nuclear Safeguards
Executive Summary (Vol. 1)

2. *(Leave blank)*

3. RECIPIENT'S ACCESSION NO.

7. AUTHOR(S)

5. DATE REPORT COMPLETED

MONTH	YEAR
November	1979

DATE REPORT ISSUED

MONTH	YEAR
February	1980

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS *(Include Zip Code)*
NRC Safeguards Program, L-97
Lawrence Livermore Laboratory
P. O. Box 808
Livermore, CA 94550

6. *(Leave blank)*

8. *(Leave blank)*

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS *(Include Zip Code)*
Dr. Robert L. Shepard
Technical Support Branch
Division of Safeguards, Fuel Cycle & Environmental Research
Office of Nuclear Regulatory Research
Washington, DC 20555

10. PROJECT/TASK/WORK UNIT NO.

11. CONTRACT NO.

13. TYPE OF REPORT
NUREG

PERIOD COVERED *(Inclusive dates)*

15. SUPPLEMENTARY NOTES

14. *(Leave blank)*

16. ABSTRACT *(200 words or less)* When setting the performance criteria for systems which safeguard special nuclear material (SNM), decision makers must consider characteristics of the adversaries who attempt to divert SNM, safeguards reponses to these attempts, costs of safeguards systems, and the consequences of diverted SNM.

This report describes an Aggregated Systems Model, which is designed to assist decision makers in integrating and evaluating these diverse factors consistently. The report summarizes the results obtained from applying the model to safeguards decision making in areas such as the hardware or procedures required, substitution of electronics for human safeguards, and overall performance criteria for safeguards systems. New performance criterion designed to measure how safeguards systems deter adversary attempts are also described.

17. KEY WORDS AND DOCUMENT ANALYSIS

17a. DESCRIPTORS

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT
Unlimited

19. SECURITY CLASS *(This report)*
Unclassified

21. NO. OF PAGES
37

20. SECURITY CLASS *(This page)*
Unclassified

22. PRICE
$

UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D. C. 20555

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. NUCLEAR REGULATORY
COMMISSION

U.S. MAIL

120555008386 2 ANRS
US NRC
SECY PUBLIC DOCUMENT ROOM
BRANCH CHIEF
WASHINGTON DC 20555

POOR ORIGINAL