IBM Center for
The Business of Government

Special Report Series

Empirically Based Intelligence Management

Using Operations Research to Improve Programmatic Decision Making

Frank B. Strickland
IBM Center for
The Business of Government

Chris Whitlock
IBM Global
Business Services

Empirically Based Intelligence Management: Using Operations Research to Improve Programmatic Decision Making

Table of Contents

Foreword

Frank B. Strickland Jr.

Kevin Green

On behalf of the IBM Center for The Business of Government, we are pleased to present this special report, *Empirically Based Intelligence Management: Using Operations Research to Inform Programmatic Decision Making*, by Chris Whitlock and Frank Strickland.

On May 14, 2012, the Office of Management and Budget (OMB) provided a memorandum to all departments and agencies on the need to use evidence throughout the Fiscal Year 2014 budget submission. The memorandum provides four pages of issues and approaches for using evidence in the development, evaluation, and management of government programs. OMB also encourages agencies to strengthen program evaluation through a dedicated senior leader, such as a chief evaluation officer reporting directly to the secretary or deputy secretary.

Reasoning and decision making with data and facts are not a result of modern management science, as a host of ancient philosophers, scientists, jurists, and others have demonstrated. However, there are some fundamental and historic changes occurring in how organizations use data to improve performance. Information and communications technologies are enabling organizations to not only gather and analyze billions of data records, but also to make sense of this data in real time. New operations technologies discern context and meaning in large volumes of data. These technologies provide organizations with meaningful performance measurements. More important, they enable the best organizations to act faster, with more accuracy and efficiency.

This opportunity also means that agency leaders and their senior staffs must equip themselves with an understanding of evaluation methods and technologies. As the OMB memorandum implies, leaders cannot delegate the evaluation function down multiple levels to specialists and still expect big impacts on program performance and management. While some of the detailed work must be performed at lower levels by specialists, senior leaders must have enough understanding of methods and technologies—as well as the program content—in order to lead. A recent series of defense evaluations illustrate the point.

In 2006, the Department of Defense found itself in a "grave and deteriorating" situation in Iraq, according to the congressionally commissioned Iraq Study Group. As part of the department's response to this crisis, the vice chairman of the Joint Chiefs of Staff—General James "Hoss" Cartwright at that time—directly oversaw a range of data-driven evaluations. Many of these evaluations focused on intelligence needs, alternative solutions to those needs, and priorities among these solutions. These evaluations directly shaped multi-billion dollar investment decisions, both for and against some capabilities. It was an intensely practical environment in which to learn about the strengths and weaknesses of evaluation methods.

Chris Whitlock and Frank Strickland are the co-founders of an evaluation approach that they and their teams applied to these defense intelligence evaluations. They not only led much of the work, but also directly presented many of the evaluations to General Cartwright. Thus, they have detailed firsthand experience with the strengths and weaknesses of evaluation methods, and have shared that understanding in this report.

Consistent with their firsthand experience, Whitlock and Strickland write from the perspective of practitioners, not theoreticians. They also understand the needs of senior decision makers, having served as senior executives in the public and private sectors. This combination of experience produces a report rich in content, but also accessible to the senior leaders and staffs who need to understand evaluation methods.

Although the authors compare evaluation methods in a defense context, the resulting strengths and weaknesses are inherent in the methods, not the programs to which they were applied. Thus, I believe that agency leaders and senior staffs will learn from this report, and be able to align the best method to the decision problem at hand. Doing so will help government leaders not only respond to the OMB memorandum, but go further in the instantiation of decision making capabilities and culture based on evidence and analytics. You can find additional information on this topic on the Center for The Business of Government's website, www.businessofgovernment.org; go to the Topic tab and select "Security, Power & Intelligence."

Frank B. Strickland, Jr.
Senior Fellow
IBM Center for The Business of Government

Kevin Green
Vice President
IBM Federal

Introduction

When the wars in Afghanistan and Iraq started our understanding of how to fully leverage unmanned aerial systems (UASs) was poor. We grossly underestimated the demands for the intelligence provided by sensors on UASs and other platforms. In the Army for example, an inadequate three-tiered UAS strategy ordered the provision of UASs at the battalion, brigade, and division/corps levels. The strategy's limited investment—one 90 minute flight time hand-held UAS with low-resolution full-motion video (FMV) at the battalion level, a six to eight hour flight time tactical (FMV only) UAS at the brigade level, and one (albeit higher quality) medium altitude endurance system at the division level—proved insufficient to aggressively target enemy insurgent networks and support other intelligence and surveillance missions. The portfolio's shortcomings, compounded by a limited series of options at the Department of Defense's disposal, created seven years of unsatisfied war-fighter demand for increased Intelligence, Surveillance, and Reconnaissance (ISR) systems support.

"This new set of realities and requirements have meant a wrenching set of changes for our military that until recently was almost completely oriented toward winning the big battles in big wars."

Secretary of Defense Robert M. Gates
Air War College
April 2008

Although demand over that seven-year period (from 2001 to 2008) was eventually addressed, the process—from understanding its capabilities, appreciating its intelligence value, and expediting the acquisitions process—was onerous, muddled, and painfully slow. Ultimately, as Secretary Gates said publicly, the military increased UASs twenty-five fold and by 2008, established the Intelligence, Surveillance, and Reconnaissance (ISR) Task Force to accelerate deployment. Incidentally, accounting for the numerous "quick reaction" capabilities fielded to deliver FMV and signals intelligence (SIGINT), the total increase would be well over twenty-five fold.

How did we underestimate the demands for this requirement so badly? After all, an estimate that misses demand by a factor of 25 (conservatively) is no "near miss." The answer is fairly simple; UAS associated technologies were new. Furthermore, military services were focused on procuring ISR systems suited for large-scale conventional force-on-force combat (characterized by Phase III of joint doctrine). Indeed, in that context, it's somewhat understandable the military's calculation was not more accurate. Nevertheless, the challenge is that targets and target behaviors in irregular warfare (central in Phase IV of joint doctrine and present in all phases) warrant more intensive ISR treatment. Strategic planners, however, overlooked experience with irregular warfare targets in the late 1980s and 1990s; experience which may have prompted a more robust debate on the size and mix of the portfolio.

The Office of the Secretary of Defense (OSD) along with the Joint Staff (and ultimately the ISR Task Force) nonetheless adjusted course and championed a series of operations research assessments to inform the debate surrounding ISR and mobilize for the "wrenching set of changes" Secretary Gates was referring to (in his April 2008 speech). The multiyear data-intensive operational assessments examined ISR systems performance across multiple mission areas in both Afghanistan and Iraq. The groundbreaking studies pioneered the application of operations research to contemporary irregular warfare. Teams processed and analyzed tens of millions of disparate records to assess ISR performance in support of specific mission areas. In addition, the assessments characterized the performance of specific capability areas such as GMTI (Ground Moving Target Indicator) or IED detection.

As the Defense Department attempts to balance looming budgetary constraints with its more traditional priorities, the time has come to take these learnings and institutionalize them in the force structure. This is not simple for two reasons. On the one hand, the Department must come to consensus on the future need. This is a complex portfolio management problem and there are competing potential approaches—modeling and simulation, subject matter expert elicitations and operational assessment being three major options. On the other hand, the classic tension between conventional and irregular-oriented capabilities forces some trade-offs. Indeed, as Secretary Gates presciently cautioned during his 2008 Air War College speech:

> I've told the Army gatherings, the lessons learned and capabilities built from Iraq and Afghanistan campaigns need to be

> institutionalized…though… if bureaucratic nature takes its course, these kinds of irregular capabilities slide into the margins.[1]

This article begins by describing a canonical operational assessment that ultimately supported major quick reaction and program of record ISR decisions and established ISR needs more reflective of the actual threats faced by the customer units. From a military operations research perspective, this study and its companion efforts stand as a body of work (complete with blemishes and garlands) used to drive leadership debates and decisions on these needs.

In fact, in the context of explaining the analytic basis for major ISR program changes ordered in 2009, then vice chairman of the Joint Chiefs of Staff, Major General James E. Cartwright (ret.) remarked:

> on the intelligence side, the work that we've done with operational research analysts out in the field on our ISR systems—not just the platforms, but how we move data and how we inform warfighters inside of the decision cycles—these analytic pieces make this as quantitative as ever I have seen in one of these budget developments.

Empirically based, data-intensive operations assessments are a break from how ISR decisions are typically informed. Traditionally, other techniques—primarily modeling/simulation and qualitative input from subject matter experts (SMEs)— were the basis for programmatic decisions made by the military services and DoD. In addition to evidence-based ISR case studies, this article explores the relative strengths and limitations of these three methods in formulating future intelligence portfolios. In conclusion, the article offers an example of how operations research is being used today in the Army G-2 to guide portfolio decision making and provoke new innovative thinking about the application of intelligence requirements.

1. Gates, Robert M. April 21, 2008. http://www.defense.gov/speeches/speech.aspx?speechid=1231.

ISR Support To Special Operations Forces—2007 Operations Assessment

Among the many operations assessments conducted from 2006 to now, a detailed assessment of ISR support to the Joint Special Operations Task Force (JSOTF) is an exemplary case that demonstrates the benefits of evidence-based operations research. In 2007, the Defense Department had to respond to a remarkable JSOTF request; a Joint Urgent Operational Need (JUON) for 30 medium altitude orbits (similar in capabilities to the US Air Force Predator) to support the mission against al Qaeda in Iraq. The JUON was remarkable because it amounted to a four-fold increase to the existing Predator fleet at a time when the Office of the Under Secretary of Defense for Intelligence (USDI) was recommending limiting the (Predator) program to 21 systems (or less). Thus, if approved, the Defense Department would be effectively granting JSOTF the equivalent of the entire fleet plus nine orbits—to accommodate one mission area. Recognizing the gravity of the decision, the USDI, along with two additional elements in the Office of the Secretary of Defense (OSD), sponsored a study of ISR support to High Value Individual (HVI) campaigns.

> *"If you have not seen the OSD HVI analysis, you need to. This is how we should be doing our work to identify and prioritize intelligence needs. We should not be doing these 'split a dollar drills.'"*
>
> Major General Mike T. Flynn
> CENTCOM J-2
> Senior Warfighter Forum
> July 2007

The resulting analysis successfully constructed a coherent assessment of the relative contributions of 13 intelligence capabilities to operational success. Starting in late January 2007, the team analyzed and associated millions of records generated by various ISR assets with data on 2,500 special operations raids against al Qaeda in Iraq. Often from relatively unstructured sources (like SharePoint, shared folders, and network diagrams) the team collected the supporting data from several disparate databases and product repositories. In addition to the bulk data-gathering phase of the study, the team also conducted direct observation tests, interviews, and focus groups.

The results were timely and coincided with programmatic budgetary decision cycles. Drafted in four months, the initial assessment arrived in time for issue development at the Pentagon (that May) and concluded before the President's Budget was finalized in December. In less than one year the team completed its exhaustive study, furnishing USDI and OSD with robust statistical insight into the performance and relative value of each capability. Isolating what was under-invested with rigorous empirical analysis, the team improved the Defense Department's understanding of what capabilities impacted performance against irregular targets.

The fundamental dataset supporting the team's analysis was a classified catalogue of daily raids conducted by Special Operations Forces (SOF). This critical repository provided results of the raid, temporal and locational data of the engagement, some indication of what intelligence cued or tipped it (e.g., human intelligence, SIGINT, FMV, etc.), and other associated data. Although intelligence cue or tip indicators proved to be of lower value in formulating our conclusions, they were, nonetheless, instructive developing hypotheses and framing other ad hoc analytics.

In addition to the operational data, the team processed tens of millions of intelligence-related messages, reports, and data. Of all the bulk data processed, the highest volume dataset was FMV telemetry; systemic observations created roughly every five seconds describing the position of the FMV platform, sensor parameters, and aim points. In addition, the team processed over 50,000 unstructured Tactical Interrogation Reports (TIRs) by creating scripts to extract all geographic coordinates, names, and other related data. Other data sources included collection management records, raid storyboards, network diagrams, source annotations, SIGINT target lists, FMV vehicle tracking files, document and media exploitation (DOMEX) records, GMTI products, and imagery.

In addition to parsing bulk data from structured and unstructured repositories, direct observation was essential to understanding the operational process known as F3EA (Find-Fix-Finish-Exploit-Analyze). First, "Find" the target—meaning identify the individuals to be pursued and understand the general location or operations area. Next, "Fix" the target under observation until a force can engage it. Third, "Finish" the tactics, techniques and procedures necessary to successfully execute the raid. Fourth, "Exploit" captured documents and media and finally, "Analyze" the data with other intelligence to fuel and repeat the cycle.

Accordingly, the team visited operational locations to observe the JSOTF and supporting elements in action. The team observed the command and tactical leadership function but concentrated on the central operations center, which handled all detainees and DOMEX. It also directly observed the daily collection management process within the JSOTF to understand the appreciable trade-offs between operational objectives and available resources.

In concert with direct observation, the team also conducted interviews to add context and balance to the quantitative analysis. Thanks to the task force's open and proactive leadership, the team was granted the freedom to probe any relevant aspect of the JSOTF's operation. This included multiple sessions with the JSOTF's commander and members of his staff. The Cryptologic Support Group as well as the HUMINT Operations Cell offered insight into the accomplishments and challenges of their respective operations. All the relevant groups provided with details (e.g. performance logs, reports) demonstrating the successes and limits of particular families of sensors. Interviews extended to domestic support as well. As an example, based locally in Washington, DC, the National Geospatial-Intelligence Agency (NGA) cell supporting the JSOTF furnished the team weekly Geographic Information Systems shapefiles included most of the relevant locational data.

Although the interviews, direct observation, and focus groups were constructive, the central element of the assessment was empirically associating the operational data with the collected bulk data. In order to accomplish this, the team recruited members with diverse skill sets, primarily professionals with extensive backgrounds in intelligence operations, quantitative analysis and advanced computing. The team created specialized computer scripts to parse the data and operational products (e.g., PowerPoint-based storyboards) to enable the analytics. In the raw product form, it would be very difficult to perform meaningful analytical work. Although the team leveraged commercial software, the tools used to analyze the data were largely limited to ArcGIS (a geographic information system), custom scripts, and conventional desktop software such as Microsoft Excel and Access.

The analysis focused on establishing spatial, temporal, and relational connections between the data and operational objectives of the raid. For example, all locational data derived from the collected intelligence datasets were geospatially plotted against each objective's location. Then, the team placed a 100-meter buffer around all locational data (objectives and intelligence) to see what sources intersected with what objectives. Since all sensors have some degree of target location error and objectives were not always individual houses (they could be larger compounds for instance) the team performed excursions to test the sensitivity of the intersections as the buffers increased from 100 to 500 meters.

Although this analysis provided a basic understanding into how "Fix-Finish" evolved, it did not provide adequate insight into the crucial "Find" phase of F3EA. For that, the team performed a range of temporal and relational tests by extracting all the names and locations produced by every detainee processed by the JSOTF. The team's subsequent analysis assessing the relationship between the locations and targets steered successive targeting. As an example, assume "Abu Muhammad" was the target captured in a particular raid. The team not only wanted to know what sources provided the location, but also wanted to know what sources identified "Abu Muhammad." This was typically something derived in SIGINT narrative reporting, SIGINT network analysis, interrogation reporting, or DOMEX. The team also was keenly interested in the temporal facets of this problem i.e., the sources that tended to lead in the identification. All the potential target names and locations produced by each detainee processed by the JSOTF. Then an analysis was performed to see how those locations and targets drove the targeting process going forward.

By May 2007, the team offered its initial position, which largely remained unchanged for the remainder of the study. In this initial assessment, intelligence capabilities were arrayed from top to bottom in tabular format indicating their relative contributions in the "Find-Fix-Finish" phases of the F3EA cycle.

The analysis revealed two striking surprises. First, the impact of FMV was critical in all phases of the process, especially in the "Find" phase and provided a substantial impact relative to other intelligence sources. Second, and perhaps even more surprising, GMTI was only a modest contributor. This was particularly unexpected because the team had recently concluded (in a separate but related assessment of the US Air

Force Predator system) the reverse would likely be true; meaning, the expectation was for strong GMTI performance relatively in tracking vehicles during the targeting process. Likewise, other "classic" ISR performers also proved to be of only modest value in irregular warfare.

Naturally, ranking contributions prompted some controversy but in performing operations assessments, this is not unexpected. Programs and their associated constituencies are characteristically disposed to only see their capabilities in the most flattering fashion, causing what we term as "constructive conflict." The team subsequently sustained a protracted period of "constructive conflict" as underlying performance issues were more closely scrutinized. For all the lower contributing capabilities, we performed a root cause analysis to chart a broader path forward. In this analysis, we blended information from direct observation, interviews, data analysis, and input from technical experts (on sensing) as well.

The team evaluated four causes for lower contribution:

1. Capacity—was performance low because we did not have enough?
2. Use—was this an issue with tactics, techniques and procedures?
3. Modification—did the capability lack some specific feature that would impact performance (e.g., downlink)?
4. Phenomenology—did the basic sensing parameters apply well for the collection capability?

The key to this analytical framework may be apparent at the surface, but is best made explicit. If a collection system is badly suited to a certain target type, no amount of money is likely to correct that deficiency. If, however, the root cause is that too few of a capability exist, or perhaps it is not being used optimally, or that it needs a modification, then the Department of Defense could take actions to improve programmatic performance.

This 2007 OSD HVI analysis ultimately developed into the foundation of our understanding of the dynamics driving intelligence performance against irregular targets. Very clearly, resolution provided by any capability emerged as a dominant theme. This was true not only with respect to spatial resolution but also temporal and relational. Continuous or near-continuous surveillance of a fleeting target was crucial to not only finding but also then fixing the target to set up action.

These dynamics were clearly illustrated, as an example, in the surprisingly positive contribution of FMV and the modest GMTI value described above. Meanwhile, the rules of engagement and basic laws of war set requirements around determining hostile intent; being confident the operators are dealing with a "bad actor" and not an innocent civilian. Identity-level resolution emerged as the grail for most of the collectors—positioning the operators to know with confidence they were actioning the right people. These basic features carried forward along a full set of subsequent complementary studies sponsored by the OSD, and later the ISR Task Force.

How Do We Approach Future ISR Needs—Three Basic Options

The OSD HVI study was but one of a number of operations assessments performed from 2006 to present under the auspices of the OSD, ISR Task Force, and Joint Staff to create a deeper, more quantitative understanding of performance. Starting with a narrow examination of the US Air Force Predator UAS in 2006 and continuing today with projects focused on operations in Afghanistan, these assessments form a body of knowledge helpful to shaping future requirements. Indeed, by April 2009 the Department of Defense had shifted significant ISR resources.

> *"Modeling and simulation has its place…for ISR portfolio assessment it is the least useful approach."*
>
> General (ret.) James E. Cartwright
> February 2012

In talking to a number of leaders from the intelligence community—including Dr. Donald M. Kerr and Gen. (ret.) Michael Hayden, both former Principal Deputy Directors of National Intelligence—it is apparent this kind of operations research-based effort is the exception, not the rule, for ISR decision making. For establishing future intelligence needs, SME prioritization efforts and modeling/simulation provide the most common alternatives (and sometimes complementary techniques) to operations assessments. Roles exist for all these methods, but our collective experience with operations assessments over the past seven years illustrates relative challenges and benefits of the three when dealing with emerging technologies, new threats, and fluid situations. SMEs are often the "go-to" method for many senior executives because they can offer immediate answers. The natural downside, however, is that their assessments are vulnerable to biases without supporting data and methods. Modeling and simulation can provide very granular outputs. Although more data-intensive, the modeling and simulation approach to enterprise intelligence performance or multisystem performance is also flawed because the generated results are often built on multiple nested and tenuous assumptions and approximations. Operations assessments measure actual system performance in concert with real missions, but identifying and collecting the right data can be challenging.

Critics may argue that operations assessments on current or recent problems offer little insight into future problems in different countries; that modeling/simulation and SMEs provide a better means to characterize and predict the future. This school of thought accepts that modeling and simulation solutions can nominally represent any chosen terrain, capability mix, and future threat forecast. Similarly, SMEs are commonly relied upon in the community to prioritize what attributes or items they believe will be important in the future. Meanwhile, operations assessments are often implicitly dismissed as “tethered” to actual events that occurred in the past or present and thus handicapped in informing the future.

Of course, no method is perfect. After all, each method requires substantial inference—none completely inoculates the decision maker from uncertainty and risk. Mixing them wisely offers a preferred path to those in the intelligence arena focused on managing the portfolio of capabilities and programs. Pushing approaches without clear perspective on the limitations can leave decision makers with the perception of greater fidelity and accuracy than is actually the case. In this context, roles exist for each of the three approaches, but great care should be taken in building an overall approach.

Modeling and Simulation (M&S)

M&S Introduction

We will start with M&S, which has powerful applications but considerable limitations in the ISR arena. To be clear, we believe M&S can make differentiated contributions, especially where the physical parameters of the problem are discrete and measureable. Where performance of various systems is known, simulation can help to establish preferred mixes in a portfolio context. The challenge in the intelligence arena is that performance is often not known or is very difficult to reasonably approximate in a model. In that instance, coarse modeling approximations and assumptions pile up underneath what may be a second decimal place output from the simulation, leaving the impression of a high-precision result despite low confidence input and assumptions.

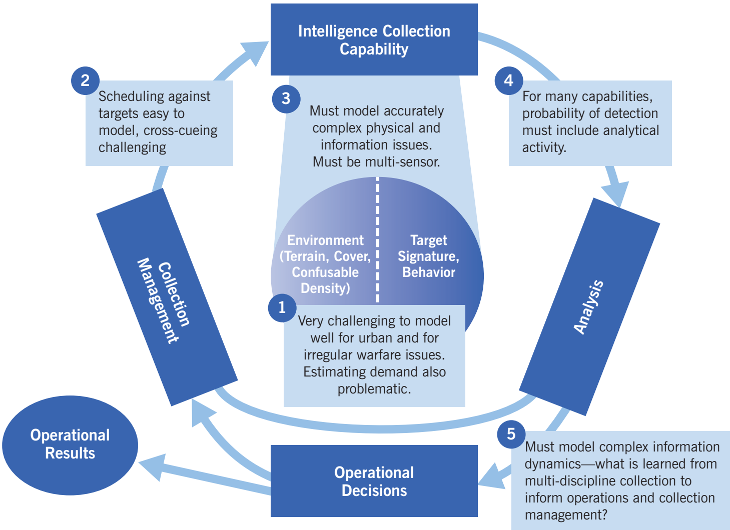
To address M&S, we will use an organizing conceptual graphic that outlines the challenges and describes the common areas where M&S efforts must approximate or estimate in order to reach portfolio mix results. Note, there is nothing inherently wrong with estimating and approximating—these are at the heart of operations research. The point is not to be deceived by very granular outputs that may mask fundamental, substantial uncertainty in crucial areas.

Figure 1 illustrates the potential problem areas.

1. First, any intelligence simulator will run against a base of targets in some simplified "environment." Estimating this demand signal and then suitably characterizing the environment and target behavior is foundational to a high-fidelity output.
2. Second, the simulator must schedule effectively against the target demand. This is the easiest element to replicate, especially as it pertains to imaging systems. The relative ease of this step can cause consumers of the result to mistake the complexities in other areas.

Figure 1: Challenges in Modeling and Simulation

Five primary areas of approximation and estimation that complicate attaining reliable intelligence portfolio mixes from modeling and simulation



3. Third, the models must capture the physics and information relationships necessary to realistically reflect what information might reasonably be collected and made available to analysts or consumers.

4. Fourth, there is sometimes a temptation to treat target detection very strictly as a physical problem, when, in fact, analysts are crucial to the equation.

5. Last, while commanders and users may sometimes directly consume a particular input, the essence of the intelligence process integrates the inputs from multiple capabilities both to sharpen the intelligence collection focus and to create insight pertinent to operational targeting and action. This multi-INT facet of the work is very challenging to replicate realistically as it encompasses a diverse range of inputs ranging from strictly parametric data to narrative reporting.

Models must necessarily simplify in all five of these areas—that is the computational reality. From a decision making perspective, it is important to understand how the uncertainties can aggregate and the implication to our perceptions of precision and suitability.

To illustrate where problems can arise in modeling ISR enterprise performance, we will explore three cases in the following section. The first illustrates the inference paradox that arises from estimating future demand with models and questions the oft-advertised belief that M&S is the technique of choice for understanding the future. The second is an actual example of a data call from a modeling team trying to tackle a portfolio planning problem regarding UASs illustrating the layered assumptions that can lie behind outputs. Finally, the third case will speak to the problems that arise from simulating "signals internals" or replicating the actual content of enemy message traffic and the simulated interpretation.

M&S Case One—Representing "Demand"

The DoD is currently grappling with size and composition of the ISR portfolio and is working to use M&S to drive the evaluations. In discussions, we often hear that operations research on current systems and capabilities is not relevant to estimating future needs because it is anchored contemporaneously or historically, while M&S can project forward into the future. So, according to this line of argument, learning on the performance of capabilities in ***Iraq, Afghanistan, Horn of Africa, etc.*** is not particularly relevant because in 2018 we will be operating "somewhere else." DoD usually articulates the "somewhere else" in the form of potential scenario contexts such as those proffered in the Defense Planning Guidance.

Circling back to the discussion of the first challenge for M&S, how do we generate the notional ISR demand signal for these future simulated environments? The simple answer in the vast majority of instances is that SMEs are asked to evaluate the scenario and forecast what intelligence they would require and the types of capabilities that should be modeled. The paradox we encounter is that the experts are drawing on experience to project this demand. What experience, you might ask? In the current environment, we are asking military members with experience in ***Iraq, Afghanistan, Horn of Africa, etc.*** to project forward regarding the need. Ironically, oftentimes those disposed to model would reject the applicability of insights derived from current operations research because it is "historical," "anchored in the current operation," or "fighting the last war," but then will go to experts and ask them to project target behavior and intelligence demands based on their personal perceptions of exactly that same experience base.

Also in this first challenge area, modelers must grapple with realistically approximating the environment, target signatures, and target behavior. In our experience, this works very well for certain kinds of environments and types of forces. The Army, as an example, has a rich legacy of sand table or tabletop war gaming (e.g. Dunn-Kempf) and computer-based simulations to help prepare commanders and leaders for conventional operations. When dealing with conventional forces, especially armored and mechanized forces, this type of modeling is very attractive. To a significant level, the behavior of these conventional forces lends itself to templating (meaning the units tend to lay out in certain arrangements and formations as characterized in the doctrine and training/exercise activity of potential adversaries). As an example, the behavior of mechanized brigades will be subject to some basic physical rules and basic tactical principles that help in creating a realistic modeling effort. Nevertheless, templating becomes substantially more challenging and much more uncertain when the fight introduces irregular warfare/asymmetric warfare components and/or urban terrain. Complex urban terrain on its own is challenging to model relative to rural or less developed terrain. Irregular warfare targets—targets dressed as civilians, using civilian vehicles and commercial communications— are tremendously challenging to model with fidelity to understand how a sensor or capability will perform.

The most fundamental ISR-oriented modeling challenge over the coming ten years may well revolve the ability to reasonably represent irregular and unconventional forces operating in urban settings. This is a major stretch today and will not likely become substantially easier over the coming years.

M&S Case Two—Data Required for Realism

To illustrate the challenges with M&S, consider an actual request for data the OSD received from a contractor group trying to use these techniques to estimate the number of UASs required to hunt insurgents in Iraq. The request (selected values represented in the table below) asked for the mean, standard deviation and minimum/maximum values for each input. The "problem category" column refers to the numbered areas in Figure 1, identifying common areas where assumptions or approximations must be made or developed. Note, only two of the items are actually facts; all others are analytical estimates. Furthermore, quite a number of the important inputs require high complexity estimates, meaning assumptions or approximations must be made in multiple categories in order to develop a value.

So, to perform an M&S based analysis of the number of UAS required to hunt insurgents in Iraq, these inputs were required to feed the application. We are making no judgment as to whether these are the right fields or a comprehensive set. Rather, we focus on the immediately obvious question: who will answer these questions and how?

Iraq covers 438,000 square kilometers (roughly), with desert areas, agricultural regions, regions with dense palm groves, several major metropolitan areas and a large number of smaller cities. The modeler wants to know the average number of confusable foot or dismounted targets per square kilometer. This is a common type of problem in attempting to realistically model and simulate ISR performance. In this instance, intelligence sources may not even exist to answer the modeler's question. Note, in the table, this input invokes assumptions in multiple problem categories.

1. First, across the totality of Iraq the population and vehicle density vary sharply. The number of potentially confusable targets will vary not only based on the terrain category (major urban, small cities, desert, agricultural, etc.) but also by the time of day.

2. Second, the resolution of the sensor will impact the number of confusable targets. Lower resolution sensors will tend to create more confusable targets, particularly in high-density areas with a good deal of intermittent masking of the targets by trees/buildings, etc.

3. Lastly, even the analysts play a role here. Some gifted intel analysts may be able to confidently and quickly exclude distracting items that to others would be a confusable target.

All this said, the simulation craves a simple value with a mean, minimum and maximum entry to represent what is very clearly a complex set of interactions.

Two other inputs listed in the table illustrate common problems in modeling. The first input is a relatively standard request for "P_detect"

Table 1: Modeling and Simulation Data Request

Input Value	Unit	Description	Problem Category	Nature of Value	Complexity in Estimate
T_Move	Hour	Mean time target moving during the day	1	Estimate	Moderate
T-Stop	Hour	Mean time target stationary during the day	1	Estimate	Moderate
T_Fly	Hour	Mean time for UAS to respond to request	3	Estimate	Moderate
T_Station	Hour	Mean time single UAS on station before rotating	3	Estimate	Moderate
T_Day	Hour	Mean time target in "Day" state	1	Estimate	Low
T_Night	Hour	Mean time target in "Night" state	1	Estimate	Low
T_Start	Hour	Mean time after beginning of "Night" before raid	Operational	Estimate	Low
T_Strike	Hour	Mean duration of a raid	Operational	Estimate	Low
Tgt_Density vehicle	#/km ^ 2	Density of confusable vehicles	1,3,4	Estimate	High
Tgt_Density foot	#/km ^ 2	Density of confusable dismounts/people	1,3,4	Estimate	High
Tgt_speed vehicle	KPH		1	Estimate	Low
Tgt_speed foot	KPH		1	Estimate	Low
Tgt_Ratio vehicle/foot		Frequency of following vehicles vs. people	2	Estimate	High
TLE_sensor	Meters	Target location error for sensor	3	Fact	Low
P_Detect	%	Probability of a sensor detecting targets	3,4	Estimate	High
UAS Speed	KPH	Mean UAS speed while imaging	3	Fact	Low
UAS Sensor Swath	Meters	Width of sensor swath	3	Estimate	High
T_Track	Hours	Mean time target can be tracked before lost	1,3,4,5	Estimate	High
T_Search	Hours	Mean time to acquire a target once imaging	1,3,4,5	Estimate	High
Tgt Cue Rate_SOF	#/Day	Rate of cues for imaging per day	2,3,5	Estimate	High
Tgt Cue Rate_Other	#/Day	Rate of cues for imaging per day	2,3,5	Estimate	High

or the "probability of the sensor detecting targets." A second interesting input requested is the "UAS Sensor Swath" representing the number of meters in the field of view at any time. In fact, these two items are not independent and are not at all represented well by point estimates.

Regarding the probability of detection, this is a function of multiple factors, including sensor resolution (inclusive of atmospheric effects), the target attributes to be detected, masking of the target, and performance of the analyst. As for "UAS Sensor Swath", realize that most of these sensors have "selectable" resolution levels or zoom-like functions in the case of FMV. "Swath" or "target width" and resolution are inherently related vice simple independent variables. But, what if the same sensor is flying at 6K, 8K, or 10K feet in altitude (all completely viable operational profiles)? How far is the aircraft track from the target location (which, when coupled with altitude, will drive slant range and the actual distance at which the sensor must perform)? How well trained is the analyst working the sensor feed? What are the atmospheric conditions? All these practical factors determine "probability of detection" which must be specified differently for vehicles versus people. Given these dynamics, notice that "aircraft altitude" and "typical stand-off distances" do not even appear as fields requested in the data call table above. We point this out simply to drive home that a simulation like the one envisioned here can output values with implied precision to multiple decimal places, when the whole endeavor has multiple, critical embedded simplifying assumptions and gross approximations.

M&S Case Three—Replicating Complex Information Environments Such as Narrative Reporting

The last M&S ISR case we will cover is an important one—modeling SIGINT, HUMINT, and DOMEX narrative reporting. These intelligence capabilities are very important in driving cues to technical sensors as well as providing insight to operational commanders on the status, potential locations, and potential intentions of adversaries as well as relationships between individuals. In real life, these things are crucial and several of these SIGINT and DOMEX functions are performed in close collaboration with interrogation and HUMINT activities. In simulation space, they are exceedingly difficult to model well, effectively get treated as assumed inputs, or do not get treated at all in the model.

Real world intelligence analysis on these types of narrative reporting puts texture onto the understanding of locations, organizations and people. The narratives are not always accurate or complete. A string of narrative reports may be analyzed over time before analysts reach a point of confidence on identifying a potential target. These types of reports and analytical processes are absolutely fundamental in shaping the intelligence results and by extension the operational results. Yet, these products represent the culmination of an obviously much messier and dynamic process than scheduling technical intelligence collection against a list of target areas of interest, which the simulator can mechanically check off to compute a success rate.

When dealing with irregular forces, the intelligence functions that generate narrative reporting are important and difficult to replicate well in modeling and simulation efforts. Classically, simulations to feed military exercises put the greatest emphasis on reporting of locational data as opposed to generating mock SIGINT and HUMINT reporting. The challenge in modeling these dynamics is unlikely to diminish while, at the same time, the importance of this type of reporting continues to rise in operations against irregular forces.

Summary View on M&S

To be clear, we support modeling in the right situation and for well-suited purposes but modeling enterprise performance of ISR is exceedingly difficult. For example communication modeling and field of view/visibility assessments can be enormously helpful to understand scheduling and collection management phenomena. In a general sense, intelligence models (despite deficiencies in realism) can be very helpful in the context of exercise support and training. But for intelligence portfolio issues, the number of assumptions and embedded estimates layered onto challenges in well representing the operating environment and targets should give decision makers pause as results are weighed.

As an illustration of valuable simulation, two of our team members performed a complex simulation and analytic activity in 2009 to assess the wide area imaging sensor visibility issues in urban terrain. Major cities caused problems for GMTI sensors but also were a locus for vehicle tracking. Wide area imaging appeared to offer a potential solution, but urban terrain masking could negate the potential value of a wider field of view. In other words, as opposed to tracking vehicles with roughly one system per target using FMV sensors, multiple targets could potentially be tracked in a bounded area by a single wide-area sensor. This analysis created a helpful understanding of the interactions of sensor grazing angle, urban elevation profiles, platform altitudes, platform orbit/speed, and target motion. Created in MATLAB and ArcGIS against a base of high-resolution urban terrain data from the war zones and using realistic track data, the resulting analysis proved very valuable for subsequent reasoning and alternatives analysis around FMV and wide area sensors.

SME-Based Approaches to Future Needs

Subject matter expert elicitations and interviews are other potentially powerful ways to inform ISR portfolio decisions but also must be used with care on complex intelligence issues. We regularly used expert input, especially from operators, for the operations research assessments from 2006-2011 to gain deeper understanding of operational details and concepts and to shape analytical hypotheses. When it comes to future needs and future utility of intelligence capabilities, using experts to create hypotheses around ISR solutions and to evaluate analysis of performance is much more powerful than relying on them for point estimates on demand or utility.

SME—Guarding Against Optimism Bias

Asking operational commanders what intelligence they find valuable can be highly useful—especially for developing analytical hypotheses—but these operator views are also prone to subconscious bias effects especially shaped by success stories. Collectively, the Defense Department holds operational commanders in high regard and rightly places value on their impressions of required assets. But the operational needs and the likely utility of intelligence capabilities are not equally easy to anticipate for commanders and their staffs.

As a test of this dynamic, we conducted a survey around a particularly high demand, but somewhat abstract ISR collection type—ground moving target indicator, GMTI. The groups surveyed included GMTI experts and analysts as well as operational leaders at various levels. When queried on the potential utility against common targets (e.g., tracking insurgent related vehicles), the operators consistently held a much more optimistic view of GMTI's performance than the GMTI experts. This contradiction illustrates a dynamic especially crucial in evaluating new capabilities or emerging applications—optimism bias.

A substantial set of material exists on this decision making dynamic. Our point is not to plumb that detail, but rather to highlight that asking operational SMEs to predict what intelligence capabilities will perform well can help generate hypotheses and context but should be treated with great care in solidifying a "validated need."

One of our lead analysts once had a very perturbed senior officer say, "if we only had GMTI available on this route, we could have found those IED emplacement teams and shut them down." The data and technical analysis would strongly suggest otherwise. Still, the senior officer developed his strongly held impressions from success stories, briefings, and flights on a GMTI aircraft and those impressions are ultimately hard to shake. In that context, it is easy to fall prey to optimism bias and operations researchers must help smoke out the realistic estimate from the operator feedback.

SME Case Example—Armed ISR

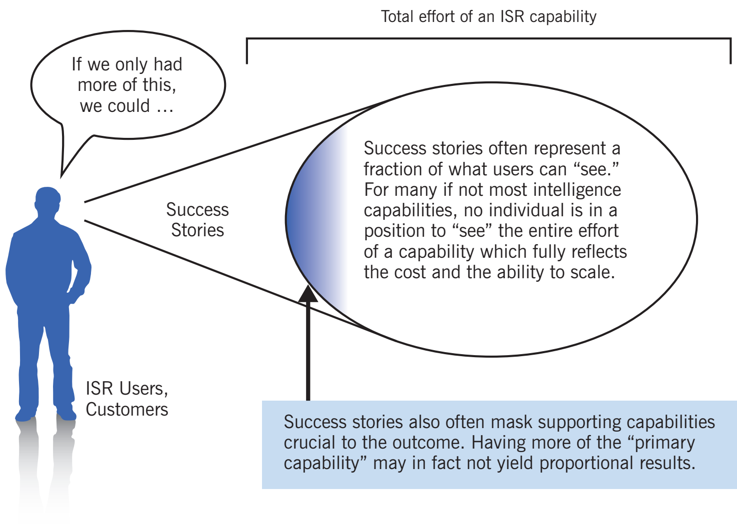
In performing ISR operations research assessments, we found success stories pertaining to intelligence capabilities can shape perceptions of the value far in excess of the actual impact.

As an example, several senior leaders believed "armed ISR" platforms were crucial to field rapidly in Iraq. There are reasons to do this, for certain, but substantial trade-offs come into play as ISR platforms are weaponized. The chief immediate trade-offs come in reductions to platform endurance but can also include substantial impacts to space, weight, and power allowances for sensor payloads.

In 2008, we were directed to evaluate a case study of the counter-indirect fire campaign that occurred in Baghdad during the summer as a strong example of why armed ISR was needed. This was an excellent hypothesis to test. In this case, a large volume of individual rockets as well as mortar attacks were fired from the northeast of Baghdad—a tough area known as Sadr City—into the International Zone, sometimes referred to as the "Green Zone". This small four by six kilometer area became the focus of intense ISR collection as well as aggressive counter-fire activity from aviation weapons teams—pairs of helicopters—and Coalition artillery elements.

Figure 2: Subject Matter Expert Potential Success Story Bias

Operators and participants can more readily "see" ISR success stories, which can lead to misperceptions of potential pay-off or ability to scale



Success stories showed (sometimes in the form of PowerPoint briefings and sometimes in the form of actual video footage) US Air Force Predators armed with HELLFIRE missiles engaging insurgent rocket and mortar teams and gave the impression this armed ISR capability was critical to shutting down the enemy's indirect fire campaign. To cut the story short and keep it unclassified, the reality was much more complicated and far less clear than suggested by the Predator success stories. The vast majority of the success we had in "finding" the launch sites was coming from other US capabilities—some designed specifically for this role. As for the Predators and other UASs with video, it turns out that even when armed ISR platforms were in place, the need to clear the heavily trafficked airspace in Baghdad to allow the UAS to fire made it easier, more advisable, and typical simply to use helicopters to engage the rocket/mortar teams. The actual number of armed ISR weapon events was actually very low, but they were disproportionately represented in success story products visible to leaders. As illustrated in Figure 2 above, users could readily see the success stories but it was much more difficult to understand the complex interactions that were preventing that from being a more routine outcome.

Our finding with respect to the Predator contribution in this case study, in our view, does not negate the broader and more complicated armed ISR issue, but it does illustrate something we found over and over with respect to "success story" slides and examples—use them to form testable hypotheses and then check to see how often the success type really occurs or could occur. Quite often, our teams found that the body of success stories represented the near full-range of successes, while not speaking at all to the hidden effort required to generate these results (the concept illustrated in the figure above). The commanders were at a disadvantage because they could not possibly see that fuller picture. In fact, most people on their staff could not replicate such a performance picture either because of data access, analytic tool limitations, analytic skill sets, etc.

SME Case Two—Be Mindful of Underlying Contextual Bias

It is difficult to forecast events and the success of new technologies. At the end of Operation Iraqi Freedom-One, the "conventional" phase of the war effort, Joint Forces Command conducted a Lessons Learned review with multiple units and organizations involved in the campaign. Some commanders clearly called for more UASs—arguing the technology was available and the units should have the benefit to inform operations. Others participating in the same operations argued that UASs would have limited applicability in stability and security operations because the platforms were too slow and the sensor fields of view too small. Both are expert views. How do we adjudicate which is correct?

SMEs implicitly bring embedded perspective that may be loaded in ways difficult to untangle. Similar to the UAS example above, it would be entirely accurate for a SME to argue, as an example, that human intelligence (HUMINT) had a limited impact on operations in Iraq. Some will bristle at this generalization, but in certain phases and in certain areas this would be definitively true. Our point is that SMEs bring a point of view based on their experience and that experience is necessarily contextualized. Experience in Iraq in 2004 is without doubt radically different from experience in 2009. In our own assessments, the impact of certain intelligence capabilities changed over time in Iraq *often based on changes in non-intelligence related factors.*

HUMINT serves as a great example of this dynamic. A key facet impacting HUMINT performance is the striking difference from 2005 to 2007 and beyond. In the former window, US forces would move into an area, conduct operations against insurgents and not remain for an extended time. As US forces departed, retaliation against individuals who collaborated with the US was common. It is quite easy to imagine the impact this had on the willingness of Iraqis to provide information. Now, changes to the OIF ground campaign strategy are well documented, most notably around the conceptualization and execution of the surge. As the strategy changed and US forces would remain in place after initially clearing out major insurgent elements, the willingness of Iraqis to work with our forces not surprisingly also changed. That change had a domino-like impact on the contribution of tactical HUMINT.

We offer this to illustrate the criticality of contextualizing a SMEs feedback. In the absence of a broader analysis, any given SME might subconsciously provide input that does not fairly reflect the likely future contributions (e.g. “HUMINT is worthless” or “HUMINT was very important to me”).

Summary Views on SME Inputs

SME inputs are crucial to understand mission context, information priorities, and even perspective on current systems. We should collectively take care in placing too much pressure on or putting too much weight on the views of operators regarding current and future systems in the absence of performance data. We must also guard against the undue impact of success stories in potentially biasing the views of operational experts. Intelligence SMEs are critical to understanding the performance of current systems and the impediments to getting further insight from extant capabilities. They are a rich source of the processing and analysis challenges. We simply must guard against overly weighting SME opinion on the performance impacts of new or modified capabilities in the absence of strong analytics.

We have found SMEs to be especially powerful in contextualizing analysis. Bring them intelligence systems performance data and let them provide caveats, reinforcing examples, and counter-points. This creates the best of both worlds. Decision makers get the richness of enterprise performance views, complemented by experiences of SMEs. These can come in the form of discrete stories and cases that illustrate the trends and dynamics in the broader set of data.

Operations Assessments—Working the Target-Capability-User Connection

Our teams conducted a series of operations research based assessments of a range of intelligence and operational issues from 2006 through to the present. Like the more detailed OSD HVI case shared earlier, each of these combined large-scale data gathering from diverse datasets along with expert interviews, direct observation, surveys, document research and some modeling. In each instance, the goal was to understand the contribution of targeted intelligence capabilities to missions or to understand the intelligence needs of particular missions. Importantly, the research highlighted the intelligence-related issues associated with reliably detecting and understanding the types of targets common in a very wide range of irregular warfare environments. Meaning, the teams dealt with the dynamics of picking select civilians out of the population, countering direct fire attacks, countering indirect fire attacks, dealing with improvised explosive device (IED) attacks, and detailing enemy irregular networks. The resulting body of knowledge is diverse, covering many different facets of intelligence and virtually all of the major military mission areas in Iraq and Afghanistan. We assert that these insights on how target types and behaviors connect to the performance of various intelligence capabilities provide a sound basis to project forward.

Consider this analogy: Based on a study of the human genome and genetics, some experts assess that readily observable differences we ascribe to race and ethnicity are explained by only six percent of the DNA content. Ninety-four percent is common. So, on the whole, the medical community does not approach humankind through the lens of race. To develop most medical solutions, researchers do not need to work each ethnic or racial group to account for differences. Largely speaking, the human body is similar across these boundaries.

We argue a similar dynamic exists when considering intelligence capabilities and decisions. By this we mean, if we can understand the performance dynamics of using FMV to track human beings in civilian clothes moving alone, moving in heavy foot traffic, moving in an urban environment, etc., then we need not study that phenomenon in every country in order to have legitimately pertinent insight. The dynamics do not likely change radically. In fact, for imaging systems, the biggest issue that affects performance will likely be the "environment": urban, rural, mixed terrain, jungle, mountains, etc. For HUMINT and DOMEX, the driver will likely be the literacy rate and the "digital index" (or degree to which computers and phones have penetrated a country's population). For SIGINT, the driver will be whether the country operates on an advanced network or whether adversaries will rely on non-telephone/internet based communications. That is a relatively small set of factors to consider to then adjust our expectations of what intelligence capabilities will play well and in what rough proportions.

Another important factor in successfully employing operations analysis is establishing a clear connection to users—either around missions or in the context of organizations. Military units are purpose built. Infantry squads, platoons, companies, etc. are designed with a range of operations and contexts in mind. From an ISR planning perspective, gaining insight into the primary operational thrusts at key unit-levels then allows a connection to the target-ISR capability data.

In the operations assessments, we were able to study both SOF and conventional forces with some significant variations in the amount of available ISR. These variations provide useful "natural experiments" that then underpin some perspective on the difference ISR can make in various missions. Further, the teams evaluated what units were requesting in the context of various types of operations. So, in contrast to the M&S approach that seeks to build a bottom-up picture of expected demand and then match a collection constellation, the operations research-based method focuses on what is needed at particular unit levels based on the types of missions they must execute.

An operations assessment team works to address the future requirement with a process that might unfold in a series of questions. For example, if divisions orchestrate clearing operations in their area of operations and brigades conduct them, we can begin to understand the parameters that influence those activities. In the simplest sense, no division clears everywhere all at once. Commanders focus, weight particular operations, and mass assets in the context of a larger operation plan or campaign plan. Several exemplar questions might follow:

- How many clearing operations might happen at once in a division?
- How many clearing operations might each Brigade Combat Team need to prepare for or execute at a time?
- When clearing operations are conducted, what are the target types and target behaviors of concern?
- What ISR capabilities map well to those target types and behaviors?

In this context, it is far less important that we precisely predict a future scenario and put simulated parameters on the environment and target. It is much more important that we establish the connection of ISR capability performance to specific target classes and behaviors and the corresponding major missions of ISR customers. In this approach, having a strong base of performance data enables analysts to infer future required capabilities with some basic assumptions regarding the different environments and the number/size of operations to be performed.

The table below reflects major assessment topics/themes over the past five years, collectively exploring virtually all active intelligence collection capabilities as well as analysis at some level. In the table, specific systems cited are those were the DoD had specific interest. Across these studies, the teams have evaluated the full-range of capabilities active in supporting ongoing operations, inclusive of SIGINT, GEOINT, HUMINT as well as processing, exploitation, and dissemination. The technical collectors included evaluation of space, aviation, and ground-based sensors. HUMINT covered not only interrogation and debriefing but also source operations to develop active assets.

Table 2: Issues Covered in Recent Operations Research

Settings	Mission Areas	Specific System Topics
• Afghanistan • Iraq • Philippine Islands • Horn of Africa	• Counter-Improved Explosive Devices • Counter Indirect Fire • Network Attack—SOF • Network Attack—Conventional • Force Protection	• USAF Predator • GMTI • Hyper-Spectral Imaging • Wide Area Imaging • U2 Optical Bar Camera • HUMINT • Space imaging systems

Summary View on Operations Assessments

Operations assessments can be a most potent way to inform portfolio decisions but must start with the difficult spadework of gathering and analyzing the data necessary to connect intelligence capability performance to target types, target behaviors, and user missions. This data is not readily available in any one place and must, in fact, be gathered using multi-faceted techniques such as interviews, direct observation, classified network crawling, and large-scale database extracts. As a consequence, the approach is not quick and generally runs six months or more to establish a clear viewpoint in a mission area or around a capability.

We now have a large body of performance-based insights to guide future requirements development, though there are areas where future research will be required. The most challenging feature of operations assessments, arguably, is keeping up with the actual performance of new or modified capabilities. The aspirations for new or modified capabilities do not always bear out, so sustaining the operations research base of knowledge becomes very important to its overall utility to portfolio decisions.

Comparing the Three Methods for Framing Future Needs/Solutions

A place exists for M&S, operations assessments, and SME inputs—all a part of operations research—but senior decision makers need to be conscious of the strengths and weaknesses of the alternative approaches and shape the decision support strategy accordingly. In our experience, senior executives play an enormous role in setting the direction for how a major issue will be analyzed. Executives would do well to grasp the trade-offs between these approaches to avoid pitfalls in the decision making process.

These three methods should be viewed as complementary approaches that can be stitched together vice competing paths to establish a portfolio view. At the highest level, we need to understand the desired attributes of intelligence and the priority missions requiring ISR support. Subject matter experts are critical to this purpose. We need modeling and simulation to gain insight on select technical parameters as well as to explore sensitivities on performance, once we understand the suitability and general numbers of capabilities revealed through operations research. Portfolio managers can leverage operations research to establish performance levels that both probe the hypotheses posed by operational experts as well as to inform modeling efforts.

Table 3: Comparison of the Three Primary Methods

Approach	Description	Strengths/Limitations
Modeling and Simulation	Computer based replications of terrain and environment with targets and intelligence collectors/analysis to establish expected future performance. Includes inputs from experts on potential intelligence demands and inputs from programs/others on expected performance.	With discrete problems and clear physical parameters (e.g. sensor visibility, transaction processing rates), this is a powerful technique. For evaluating multisystem performance, many layered assumptions and estimates can render this approach an inscrutable "black box." Very difficult to accurately capture irregular warfare and urban environments with rigor.
Operations Assessments	Analysis of actual intelligence capabilities performance in the context of operations, major policy decisions, intelligence gaps. Includes interview, direct observation, bulk data gathering and analysis.	This method works from outcomes and results backwards, providing insight into use of and value of capabilities. Framing the analysis is critical and can be derailing. Data access and access to operational participants important—manageable but requires purposeful engagement.
Subject Matter Expert Elicitations and Prioritizations	Interview and focus group-based interactions to understand operator perspective on intelligence needs, solutions, priorities. Decision making techniques that have operators weight alternative scenarios and solutions.	Potent method for framing hypotheses and to understand operational context. Short of direct observation, this is best for understanding. When shifting to solutions and priorities, beware subconscious optimism and other biases. These techniques can easily ask for answers well exceeding "expertise" base.

Using Operations Assessments to Conceptualize the Army Integrated Sensor Coverage Area Framework

The Army ISR Challenge

The Army needed a more effective approach to understanding the ISR need and thinking about ISR employment. Through much of the last decade, anticipated demand for ISR assets in combat theaters was calculated imprecisely because estimates were based on projected "hours" of use and expected "orbits" of capacity vice functional unit-based requirements suited to specific mission requirements. This "hours" and "orbits" method was problematic because:

> *"The military services are inconsistent and imprecise in defining [UAS] requirements…the Army should begin using the Integrated Sensor Coverage Area concept immediately. This is a much more conducive approach to cost-benefit analysis."*
>
> House Permanent Select Committee on Intelligence ISR Program Review February 2012

a) "Hours" were aggregate and obscured distinct but concurrent coverage requirements. A single 10 hour requirement, for example, is different than two five-hour requirements (over that same period).

b) "Orbit" endurance and available collection time vary by platform. The appreciable difference, for example, between the duration of an Army Enhanced Medium Altitude Reconnaissance and Surveillance System "orbit" and USAF Predator "orbit" (six versus 20 hours respectively) is illustrative of the fundamental problem employing "orbits" as a unit of measurement.

c) "Orbits" and "hours" mask crucial multi-intelligence collection needs— instead often being articulated for individual types of sensing (e.g. Full Motion Video). As the team's examination of ISR performance in combat theaters conclusively demonstrated, no single intelligence discipline or capability stands alone.

d) In terms of direct ISR support, “hours” and “orbits” are not linked to unit-level mission requirements. In our recent past, Brigade Combat Teams, Divisions, and Corps had a very modest organic ISR capability and then once in conflict requests for additional ISR are presumably to be satisfied by joint assets or Quick Reaction Capabilities. Congressional staffs, however, characterized the resulting ISR requests, needs as seemingly “insatiable.” One of the root issues is that for the forces conducting operations, beginning at the battalion and brigade level, we have both underestimated the enduring needs and failed to provide the right solutions.

A New ISR Concept Emerges

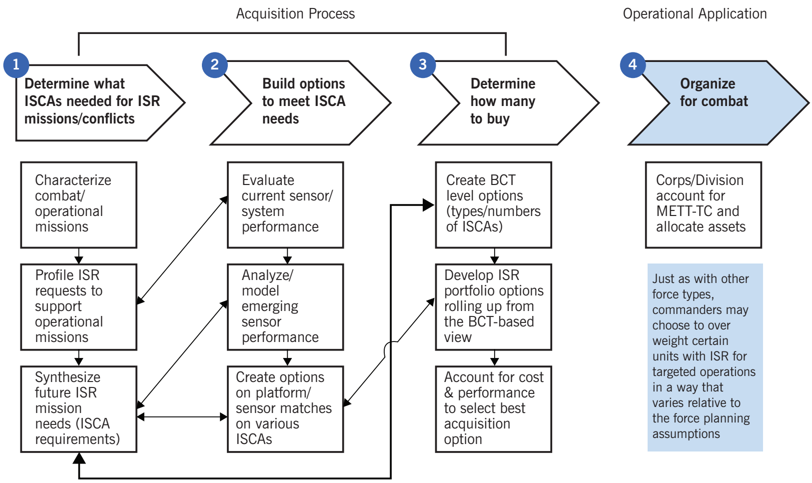
The Army developed an innovative solution—leveraging operations assessments—building around several different types of multi-intelligence, multi-sensor ISR needs apparent at the BCT-level. In contrast to the “hours” and “orbits” method, this concept—referred to as Integrated Sensor Coverage Area (ISCA)—starts with integrated ISR functionality, based on the observed performance in the war zones. Building on this basic functionality (i.e., the capabilities the requestor needs to support the mission) technical parameters are framed for each ISCA type using insights derived from operations research. Then solutions and portfolio options are developed to address identified requirements.

Operations assessment-based insights drove ISCA conceptualization process in a way difficult to replicate using Modeling and Simulation and with rigor not possible with a SME-based approach alone. Modeling and Simulation, as discussed, can be very helpful in evaluating options or potential performance. But, the analyst needs actual options to evaluate. Operation assessments provided empirical insight on the performance of various sensors and capabilities necessary to enable conceptualization of new combinations and new approaches. Expert inputs were important to the operations assessments to focus the analysis of capabilities. They were not treated as the “gold standard” with respect to need, but rather as a guidepost for the analysis approach. In this respect, the development of the Army’s ISCA concept and the elaboration of the associated needs illustrate the relative role of the methods in conceptualization. In the evaluation of options phase, modeling and simulation clearly can play a more significant role, if the challenges can be overcome to achieve realistic outputs. Similarly, expert inputs can be crucial in the evaluation stage to gain perspective on priorities and a sense for what operators might value most in future settings. Either way, the development of the ISCA illustrates the use of operations assessments to create a future framework for the Army and potentially for the DoD in considering ISR solutions.

Figure 3 outlines the major steps and activities in developing the ISCA concept and speaks to how operations assessments were used in the process:

Figure 3: Portfolio Development Approach

Initially, the problem was viewed in four phases with the emphasis being on the first three



The first phase in the figure above was arguably the most important in that it moved away from aggregating hours of coverage to identifying three ISR functions in the form of the ISCAs. Working with the Army G-2, the team reflected back on the operations assessments in Afghanistan and Iraq, as well as some studies performed in other operations theaters. Starting with a focus on the primary missions performed by the combat forces, four principle missions emerged as central to framing the ISR requirements needs (network attack, interdiction, force protection, and population protection). These missions were routinely conducted in the context of the "Shape-Clear-Hold-Build" cycle and represented target sets and behaviors that would be present in a number of other irregular conflict settings. The team looked back across ISR requests to understand the major categories and themes. For example, when conducting vehicle following, how large is the required coverage area? This type of work is a combination of the analyzing ISR requests as well as synthesizing insights from operational assessments of actual ISR performance. This included namely identifying the core ISR capabilities and how they mixed (with other capabilities) in different contexts. These observations on the needs were subsequently synthesized in the last section portion of the first phase to create three ISCAs—Persistent Area Assessment (PAA), Situation Development (SID), and Mission Overwatch (MO).

Next, phase two evaluated and analyzed performance data on current and emerging ISR sensors and platform sensors with respect to the three types of ISCAs. The team probed interactions between different types of sensors from a performance perspective as well as platform sensor interactions. As an example, on the imaging front, attention was focused on resolution required to effectively sense the behaviors and observables as well as the typical number of targets per ISCA. With respect to FMV, the team assessed the resolutions typically chosen when operators had a selection using metadata from actual sensors. For wide area systems, the team evaluated the causes of tracking failures in wide area imaging systems to gain insight on resolution and

Army Military Intelligence Rebalance

This is a comprehensive strategy led by Lieutenant General Richard P. Zahner, Deputy Chief of Staff, Army G-2, to realign substantial military intelligence resources and programs to improve anticipated support in future conflicts.

LTG Zahner is a voracious consumer of operations research and used this material as the basis for major decisions of the ISR Task Force as well as in his subsequent leadership as the Army G-2.

LTG Zahner used his substantial direct experience and expertise to lead this strategy, then leveraged operations research and R&D to guide the framing of requirements and solutions. Many of the elements of this strategy have already been acted on by the Army, while major decisions are pending for the aerial layer components.

Unclassified briefings of the Rebalance Strategy can be found at: http://www.dami.army.pentagon.mil/site/G2%20Vision/pod/IWS%20Brief.pdf area coverage dynamics. From a platform perspective, some micro modeling was performed to understand the impact of orbit types on sensing potential with special focus on understanding the notional differences between rotary- and fixed_wing UAS in tracking targets with narrow field of view sensors. The insights were synthesized to create a set of platform-sensor combinations and optional approaches for each type of ISCA. To illustrate, at this point the team could envision two potential paths to address the PAA ISCA needs—using several Army Gray Eagle's specially equipped or using a new Long-Endurance Multi-purpose Vehicle (LEMV).

In phase three, the team developed a mix of options for the Army to determine the optimal number of ISCA systems to buy to address each respective requirement. Careful to balance capability with affordability, the team focused more on routine operational tasks at the battalion and brigade level than trying to create a granular bottom-up model based on wholesale demand. For example, two proposed PAA/SID/MO mixes per Brigade Combat Team (BCT) were 2/3/.25 and 1/3/.25. In other words, representative options per BCT would include one or two PAA ISCAs, three SID ISCAs (one per battalion as a planning factor), and .25 MO ISCA. For each ISCA type, different system combinations could address requirements at different cost levels. Overall, phase three put the Army G-2 in a position to articulate a preferred level of ISR support in the form of ISCAs per BCT as well as a point of view on the relative costs of the various system options.

Operations Assessments Role in Developing the PAA ISCA

Recently, the only ISR capability able to monitor an area continuously for a BCT was the JSTARS, assuming such coverage would be available at all in the BCT area of operations. Each BCT is configured with a ground station to receive the JSTARS GMTI sensor data (or dots) that represent ground movement. Two challenges emerged in the irregular warfare environment. On the one hand, the ground forces hand a large appetite for additional wide area surveillance, well beyond what was available in the form of current GMTI and wide area imaging systems. On the other, the current GMTI capabilities were really of insufficient resolution to work the problem effectively. To compound the problem, GMTI typically required additional high-resolution ISR (already in short supply with many competing priorities) to follow up on the relatively low-confidence leads the system would generate. As an example, this Air Force Magazine quotation from a 2009 article captures the insufficient resolution and collection management challenge in the current scheme:

> For the time being, the E-8 [JSTARS] won't tell you what model of Toyota 4Runner the insurgents are driving, but it will tell you that there's unusual movement in specific places… [JSTARS can then] hand off that information for closer inspection by Predator surveillance drones or manned aircraft.[2]

Even this quote is generous, in that a JSTARS cannot distinguish reliably between trucks or cars. The GMTI operator is presented with a "dot" on screen that represents detected movement of something. Predominately, the detections are of vehicle movement. In the original conception, the JSTARS was intended to reliably detect and track the movement of armored formations to set up deep attacks. Now the system is called upon for traffic pattern analysis and even for tracking. But

2. Grant, Rebecca, "JSTARS Wars," *Air Force Magazine,* November 2009

the limitations are substantial in resolution. This, in turn, typically prompts a response for yet another ISR platform to confirm, deny, or follow-up on suspicious activity. Often, these necessary additional ISR systems are also in short supply. Taken together, in the current instantiation, GMTI has not been particularly effective and not for want of sincere effort.

Conceptualizing the PAA ISCA—Phase One

In phase one of Figure 3 (page 38) in creating the PAA ISCA concept, the team was not simply trying to address a shortfall in the aggregate number of GMTI collection hours, but rather to capture the more meaningful multi-discipline intelligence problem for area coverage supporting a BCT. The integration of the multi-discipline sensing is crucial because seldom is one type sufficient especially regarding irregular warfare targets. The conceptual design questions are relatively straightforward: how big an area is relevant and what types of sensing might be combined usefully? The challenging element in this arena is to determine the area-resolution combination that is effective both operationally and from an ISR perspective. It is simple to generate an area "requirement" that far exceeds the ability to effectively detect and track the targets. This is a deficiency with existing area collectors.

The team analyzed tens of thousands of actual high-resolution target tracking activities performed with narrow field of view sensors to understand the need. The issue was to characterize how far targets typically traveled, how many stops they made, and what other activities occurred in the context of the tracking that would be highly relevant from an ISR perspective. Using actual high-resolution tracking data from multiple areas of operations, the team derived insights to help bound the size the required area and the types of sensing to be combined.

Determining PAA Sensor/Platform Combinations/ Options—Phase Two

The team then reviewed operations assessment performance data on current wide area collection systems to understand performance drivers and limitations. This analysis not only included radar systems like the JSTARS but also newer quick reaction capabilities like the Constant

Hawk or Angel Fire, which used imaging technologies to monitor larger areas. Last, emerging sensors (e.g., the VADER GMTI or ARGUS-IS) were evaluated in performance specifications and test data to understand potential capability that might realistically be included in future systems. Finally, the needs (area, number of targets, types of collection, etc.) were coupled with sensors and platforms not only to create an ISCA with parameters but also some sensor/platform combinations that would satisfy the need. Interestingly, the PAA ISCA does not match up well to any existing system (sensor/platform combination), so a new option had to be built out around an airship-based concept to account for the required sensor combinations. Using a long endurance airship as a platform, an array of mutually supporting sensors can be hosted both to provide the necessary ISR coverage as well as eliminating the substantial collection management challenges in the current approach.

Determining the Number of PAA ISCA per BCT—Phase Three

The PAA ISCA did not emerge from a specific user solution request or simple aggregation of hours of requested collection. It is the result of operations assessment that begins with consideration of the needs of the units—BCT and its battalions—testing the boundaries and the needs using actual operational data and requests. Rather than perpetuating single intelligence discipline, independent planning, the PAA ISCA accounts for all the aerial ISR needs associated with the BCTs area collection problem. While founded on the irregular warfare target observations, the solution migrates relatively gracefully into the needs stemming from the "conventional" phase of a conflict.

In determining the number of PAA ISCAs required, the focus was far less on a number of collection hours than on the missions and operational environment of a BCT. While a commander may desire complete situational awareness and complete coverage of the battlespace all the time, the practical reality is that there are choices and trade-offs. A BCT on the attack, in defense, or in a wide area security mission (as is common in irregular warfare) will likely have a main effort, a priority avenue of approach, or some nagging problematic area. A BCT has three battalions and ten or eleven companies to put into an offensive posture—each one cannot be supported with a PAA ISCA. Given that there are cost boundaries and limits on the number of intelligence analysts that can be mustered, it was relatively simple to set one or two

PAA ISCAs as the desirable planning factor per BCT. This number of PAA ISCAs would provide the commander an ability to "soak" an area for days prior to major operations (battalion-sized) or to conduct sustained surveillance over a problematic area (i.e. repeated direct fire, indirect fire, or IED attacks) to support network attack and force/population protection.

The last portion phase three deals with affordability and cost, evaluating various options for delivering the targeted number of PAA ISCAs. In the Army context, each ISCA would require either one airship-based system or multiple Gray Eagles. Initial quick-look costing was conducted using data available in the Army G-2, with a more formal iteration performed later with the Army cost estimating team. The costing effort evaluated acquisition, O&M and personnel related costs of each option to provide the Army a perspective on the path forward.

Conclusion

The Integrated Sensor Coverage Area (ISCA) concept represents an innovative shift from the conventional imprecise "hours" method of anticipating demand for ISR support to a more economical unit-based framework tailored to complement mission requirements and optimize resources. Derived from operations research, the parameters of each ISCA were formulated based on insight from actual operations and the performance of various ISR capabilities in the conduct of those operations. Thus, military commanders can now expect a basic level of ISR support based on force level deployments and functional requirements vice "hours" requested. Indeed, each ISCA is rendered as a 24-hour block with associated sensing and target requirements as well as SIGINT and imagery-based (radar inclusive) capabilities. ISCA purposefully links multi-sensor requirements to operational force levels and fixed duration times. Accordingly, force planners and collection managers no longer have to juggle capabilities with time.

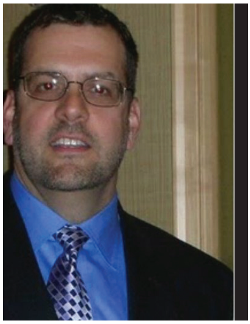
About the Authors

Frank B. Strickland, Jr. is a Senior Fellow with the IBM Center for The Business of Government and a Partner in IBM's Global Business Services.

Prior to joining IBM, Mr. Strickland co-founded Edge Consulting, a consulting firm that achieved national recognition for pioneering work in the application of operations research methods and IT to quantify the value of intelligence. He helped lead Edge Consulting from a start-up to significant annual growth, culminating in its acquisition by National Interest Security Company.

Mr. Strickland was a career intelligence officer with 24 years experience in the Central Intelligence Agency's Senior Intelligence Service and the U.S. Marine Corps, where he led programs focused on developing innovative solutions and methodologies to measure and analyze mission performance. In recognition of his accomplishments, the CIA Director awarded him with the National Intelligence Medal of Achievement. Mr. Strickland also received the National Reconnaissance Office's Medals of Distinguished and Superior Service.

Mr. Strickland is the co-creator of "Edge Methods," a unique blend of consulting, scientific methods, and IT used to assess the value of information from empirical data. Edge Methods has been used to advise national security principals and commanders on the optimal use of billions of dollars of operational and fiscal intelligence resources. He is a recognized teacher, public speaker, and published author. He holds a BA in Business Management, MS in Technology Management, and the CIO University's Certificate in Federal Executive Competencies.



Chris Whitlock is a Partner in IBM's Global Business Services.

Chris Whitlock has worked defense and national security issues for the past 30 years. For the last 20, he focused primarily on strategy consulting on intelligence issues from an analytic perspective. He co-founded and was the CEO of Edge Consulting, which applied empirical methods and management consulting techniques to advise on major programmatic issues confronting DoD and the Intelligence Community. He holds a B.A. in History (Mississippi), an M.A. in National Security (Georgetown) and an M.B.A. (George Mason).

Chris was trained as an Infantry officer in the United States Army and subsequently became a military analyst with the CIA. He is an expert in technical forms of sensing, having spent most of the past 20 years working projects to improve performance for imagery and signals intelligence systems. He has worked target problems in a variety of forms and countries, including Panama, Colombia, El Salvador, Bosnia, the former Soviet Union, Korea, Iraq, Iran, and Afghanistan. He has led and worked on projects cutting across a range of mission areas including counter-insurgency, theater missile defense, counter-fire, suppression of enemy air defense, combat search and rescue, and combined arms maneuver. With respect to methods, Chris has applied multi-attribute utility theory, modeling, and various empirical approaches including the development of a Warfighting Applications Research approach and what came to be known as "Edge Methods."

IBM Center for
The Business of Government

About the IBM Center for The Business of Government

Through research stipends and events, the IBM Center for The Business of Government stimulates research and facilitates discussion of new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

About IBM Global Business Services

With consultants and professional staff in more than 160 countries globally, IBM Global Business Services is the world's largest consulting services organization. IBM Global Business Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build, and run those solutions in a way that delivers bottom-line value. To learn more visit: ibm.com

For more information:
Jonathan D. Breul
Executive Director
IBM Center for The Business of Government
600 14th Street NW
Second Floor
Washington, DC 20005
202-551-9342
website: www.businessofgovernment.org
e-mail: businessofgovernment@us.ibm.com

Stay connected with the IBM Center on:

or, send us your name and e-mail to receive our newsletters.