Deutscher Luft- und Raumfahrtkongress 2016
DocumentID: 420077

CONSIDERATION OF SURVEILLANCE SENSOR CAPABILITIES WITHIN THE HOLISTIC EVALUATION OF AERIAL PLATFORMS

S. Morawietz, M. Strohal, P. Stütz
Institute of Flight Systems, University of the Bundeswehr Munich, Neubiberg, Germany

Abstract

In this work a method for supporting a holistic evaluation process through the derivation of a descriptive metric structure in combination with relevant decision parameters is presented and implemented on the example of surveillance sensor capabilities. The approach is based on the *House of Quality* (HoQ), which was modified to be used for systems analysis/evaluation. Thereof a new method named *House of Metrics* (HoM) was derived to compile and manage all relevant parameters needed for today´s complex evaluation tasks. The implementation of the methodical approach is demonstrated on a simplified scenario to show the contribution of a surveillance mission sensor sub-system to the mission-related performance evaluation of (un)manned aerial platforms. Thus a hierarchy of holistic metrics in combination with the corresponding weighting vectors was developed. Furthermore a modeling and simulation environment is presented to quantify e*lementary criteria* for, inter alia, environmental perception as a prerequisite for the evaluation process to assess two considered systems. For system evaluation a method of *multicriteria decision analysis* is used to rank the considered systems and to verify the robustness of the results for deriving a surveillance and reconnaissance mission through a sensitivity analysis.

1. INTRODUCTION

Due to the introduction of new technologies in connection with complex mission scenarios the exclusive contemplation of flight performance for the evaluation of aircraft concepts is insufficient [1]. In the past, aircraft evaluation has been characterized by an emphasis for optimum flight performance which was characterized by attributes such as speed, payload or endurance which may lead to irreversible decisions with enormous consequences through incomplete evaluation results.

This particularly is inappropriate for the majority of today´s aerial reconnaissance and surveillance platforms, especially UAVs using specific mission sensors [2]. In this respect considerable technological advances in the last few years enhanced performance characteristics of imaging payloads while reducing power, size and weight characteristics. Therefore considering actual sensor surveillance capabilities for the overall analysis of current aerial platforms has become not optional but mandatory and emphasis has shifted from performance to overall system effectiveness as a key measure of merit for the aerial system [3]. Withal onboard environmental perception capability is a key metric for the overall system effectiveness of military aerial systems. Thereby the complexity of the evaluation task increases which requires a structured and interdisciplinary approach. Consequently, holistic mission-based approaches need to be developed, which consider the aircraft as a product of individual subsystems (*System of Systems*) and derive a descriptive metric structure as well as additional relevant parameters for decision making therefrom. In this work, we present a method to consider the complexity of such evaluation task and to derive a holistic metric structure in combination with relevant evaluation parameters using the example of a surveillance and reconnaissance (S&R) mission. As a use case, we consider an S&R mission being conducted by an aerial platform with different mission sensors to show the impact on overall system effectiveness within the scope of a related scenario described later at chapter 5.1.

2. RELATED WORK

Great efforts have been made in military *Operations Research* (OR) to evaluate weapon systems in interconnected relationships to achieve campaign objectives. The output parameter of these models are usually metrics used to indicate the quality of multiple systems called *Measures of Effectiveness* (MoE) [4]. Typical examples of campaign level MoEs are the relation of deployed to damaged sorties or the time to get combat superiority. Military decision makers use these models to answer questions concerning the probability of success and required resources. While these models deal with the evaluation of the overall campaign scenario comprising ground, air and maritime entities [5], this work is concerned with the evaluation of an single entity (aerial platform). Nevertheless the quantification of *System Effectiveness* (SE) can be adopted from campaign to mission level to assess the considered system including the corresponding subsystems. In this context [3] noted that the concept of SE was defined primarily as functions of the “*-ilities*”: e.g. reliability, availability or survivability. More attributes and their definitions as well as the historical background of SE models can be found in [6]. As such, the system effectiveness concept was applied from OR to the aerial platform and further to corresponding subsystems (e.g. mission sensor), depending on the evaluation task. Therefore the criteria structure of the evaluation catalogue in this work was separated into an ability (containing the “*-ilities*”) and criteria block depending on a related mission block (see chapter 4.1).

Numerous methods can be found in literature to assess and trade off the impact of various technologies on the overall system effectiveness during early design phases. For example, firstly the method of “*Technology Impact Forecasting*” [7] using physical design parameters as standard level metrics such as aspect and tapper ratio as well as maximum thrust. Secondly the method for performing early technology tradeoff and design studies which was presented in [8] dealing with the direct mapping between the *Quality Function Deployment* (QFD) methodology and standard design space exploration techniques for improving the trade studies and design precisions. Therefore the *House of Quality* (HoQ) as a standard QFD tool is used to map the customer requirements to a set of engineering characteristics with the ambition to understand the overall relationships. Furthermore a set of surrogate models as *response surface equations* were defined to explore the design space through mathematical functions [9].

3. FUNDAMENTALS

A overall concept for the evaluation process is already provided in [10]. The concept consists of a domain-independent macro-concept which includes a process model based on the fundus of systems engineering and an integrated domain-specific micro-concept which deals with the compilation of a holistic evaluation structure to assess aerial platforms related to the specified mission(s). The simplified process model of the macro-concept is shown in Figure 1.

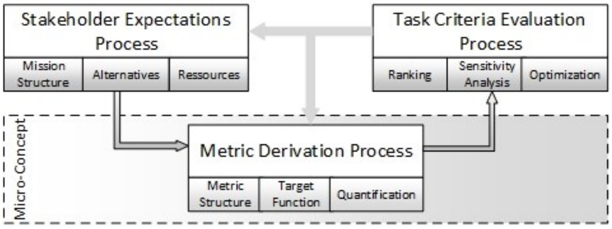

Figure 1: Process model for the evaluation framework of the macro concept from [10]

The *Stakeholder Expectations Process* provides the initial requirements for the decision and evaluation process. These are for example the considered missions (weighted structure) and alternatives (platforms with different mission sensors) as well as contributing resources like subject matter experts (SMEs). The subsequent *Metric Derivation Process* provides a methodical approach for deriving a mission-related evaluation structure consisting of weighted criteria in combination with target functions and quantification approaches. The method for the *Metric Derivation Process* is based on the *House of Quality* (HoQ) which has to be broadened to meet the complexity of the concept. This work now details the derivation of this method in chapter 4. The *Task Criteria Evaluation Process* uses methods of *multi criteria decision theory* providing *"a common language and approach that removes decision making from the realm of personal preference"* [11].

The HoQ (formerly known as quality table) is the most recognized and widely used matrix format to capture a number of essential issues for *Quality Function Deployment*. QFD is a proven quality engineering technique that has been credited significant reductions in resource allocation, as well as increases in product reliability when applied correctly in development process [12]. It helps to identify the most important engineering characteristics which fulfill the customer needs. The traditional HoQ consists of several parts illustrated in Figure 2. The customer requirements ("whats") are given in part 'A'. After having determined which requirements are most important (part 'AA') to the customer, their needs have to be translated into design goals ("hows") which are listed in 'B'. The corresponding direction of improvement called targeting function is listed in 'BBB' (e.g. maximize means "bigger is better", minimize means "smaller is better"). To populate the room (part 'C') the requirements and engineering characteristics are contrasted and the corresponding strength of relation or dependability is listed. Thereby, the relations are specified on different scales which were mostly converted from verbal codes and indicating symbols to nonlinear numerical scales (e.g. "strong" = 9, "middle" = 3, "weak" = 1). The choice of the scale is variable and dependent on the use case so that several types of scale classification exist. Mostly the chosen scale is nonlinear so that a strong relationship counts more than a weak one. For each engineering characteristic its importance (weight_{eng}) is calculated at the basement in part 'BB' by using Equation (1) through multiplying the technical importance of each requirement (weight_{requ}) and the relation score between the engineering characteristic and the requirement. Subsequently the columns are summed up [12].

(1) $$\text{weight}_{eng} = \sum_{i=1}^{n} \text{relation score} \times \text{weight}_{requ}$$

The correlation matrix is established to determine the interrelationships. The correlations between the "whats" are given at the 'dooryard' in part 'D' and respectively between the "hows" at the 'roof' in part 'E' on an underlying scale. According to [8] these correlations can be used to help identify where critical design tradeoffs may be expected to occur. A hierarchical linking of several HoQ is a common practice in QFD. So the "hows" with their calculated weightings become the "whats" of the next HoQ until the desired level of detail is achieved. Summing up, the *House of Quality* is used as a worksheet to establish means for planning and communication between people with different problems and responsibilities (costumers and suppliers). Hence the HoQ provides a promising foundation to manage the complexity of a multidisciplinary system analysis task.

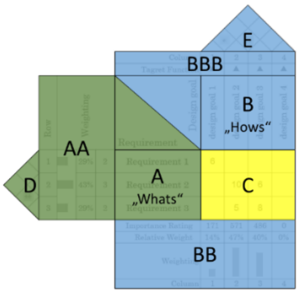

Figure 2: simplified HoQ structure

4. CONCEPT

To consider sensor based surveillance capabilities the already mentioned micro-concept is used. The process model of the overall macro-concept was shortly discussed in the previous chapter. Subsequently, the main focus is now on the *Metric Derivation Process* which pursues the approach of *System Effectiveness* and uses a block structure in combination with a derivative of the *House of Quality* to identify descriptive metrics and essential decision functions.

4.1. The Top-Down block structure

As already mentioned in chapter 2 the concept of *System Effectiveness* was defined primarily as function of the "*-ilities*". Therefore the overall *System Performance Potential* (SPP) comprising of the *System Effectiveness* and *System Efficiency* on the mission block is delineated by an ability and criteria block. The generated block structure is shown in Figure 3. The ability block is attached to the mission block and contains the necessary capabilities of a system to meet the related mission requirements. Thereby it comprises several capability components which were primary developed by the planning staff of the German Federal Armed Forces and adapted in this work for the consideration of surveillance sensor capabilities. Such capability components for the considered S&R mission are for example the availability, operational capability, survivability and environmental perception capability. With regard to *System Efficiency* one component inside the ability block is the arising cost which could be listed separately or allocated inside the abovementioned capability components to define *System Efficiency* [10]. Underneath this block a further block joins containing the criteria for describing the mentioned capabilities. This block is called criteria block and holds the standard metrics for the evaluation in different levels of detail. The block structure is shown in Figure 3.

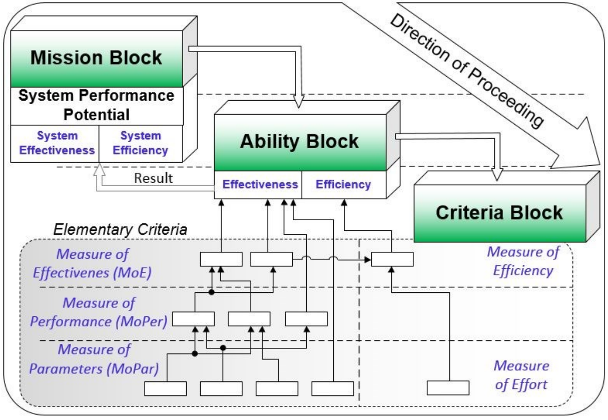

Figure 3: Block structure and e*lementary criteria* levels inside the derivation process

Mission block:

The mission bock contains the processing evaluation task in terms of weighted mission components. Therefore an extensive mission structure was established which represents an overview over nowadays civil and military applications of aerial platforms. The *Stakeholder Expectations Process* (see Figure 1) is used to derive the evaluation task in terms of weighted missions. For the following explanations the mission block should not be considered in detail. In the toy scenario the mission block consists only of one submission which is the simplest use case that may arise.

Ability block:

Based on the mission block, quantifiers have to be specified to indicate the system performance to identify the variant with the best eligibility to fulfill the evaluation task (S&R mission). These quantifiers take the form of weighted combinations of several other capabilities. For example the before mentioned capabilities for a S&R mission can be decomposed further to get an approach for quantification which means to get their intensity through making them predictable. For example, the capability of *environmental perception* consists of several capabilities on the next level of detail like the capability to generate data or information, the capability to transfer or provide data in combination with the real-time capability and the interoperability with other systems or services.

Because the capabilities inside the ability block are formulated in abstract terms, a further decomposition is necessary to quantify the capabilities by summing up the decomposed predictable results. Thereby the top level criteria (*Measure of Effectiveness/Efficiency*) or every underlying criterion (e.g. *Measure of Performance/ Parameters*) out of the criteria block are attached to the leaves of the ability block. For example the *environmental perception* capability on the second level consists among others of the probability of detection, recognition and identification (DRI) as well as the target leakage and nomination rate. Such can be quantified only at this decomposed level inside the criteria block by *Modelling and Simulation Environments* (MSEs). Thereby the above-mentioned children of the *environmental perception* capability are listed at the criteria block through their possibility to be quantified. Therefore the ability block serves as a connecting block between the mission block and the metrics inside the criteria block to map the mission as complete as possible. The usage of capabilities encourages the structured approach to derive mission-related metrics through the attachment of criteria structures on the capabilities leaves along the principle of a construction kit.

Criteria block:

At this block the detailed composition of the abstract capabilities is primarily shown through significant related criteria. The top level is divided into the *Measure of Effectiveness* (related solely to utility) and *Measure of Efficiency* (related to utility and effort). Only just at this stage the criteria is quantifiable and their intensity could be calculated through MSEs. Depending on the complexity and the availability of the MSE the top level could be decomposed in further levels of detail. For example with a complex simulation environment it is feasible to calculate top level criteria (*Measure of Effectiveness*) like a detection rate, target leakage rate or target nomination rate. Such simulations have to consider the hardware as well as the software capabilities of surveillance sensor systems including capabilities of resource management, multisensory data fusion and environmental conditions.

The leaves of the criteria structure are called *elementary criteria* which should be located on the highest aggregated criteria level that could be modelled and simulated. It can be read as follows, the greater/smaller the level (*Measure of Effectiveness/Performance/Parameters*), the greater/ smaller the complexity of the MSEs, the grater/smaller the accuracy of the result. The endeavor should be to locate the *elementary criteria* on the highest level inside the criteria block, because every decomposition inside the criteria block to a deeper level leads to increased deviation at the result accuracy through the loss of information. In case of lacking complex MSEs the criteria has to be decomposed to the *Measure of Performance* where the MSE could be more simplified and specified to a particular criteria. On the lowest level (*Measure of Parameters*) only parameters were considered which mostly do not require any MSE. For example the detection rate can be decomposed to standard sensor parameters like the *signal to noise ratio*, the *spatial* and *spectral resolution* or the *focal length* which result in the *ground sample distance* by using the sensor object distance, and the *contrast* of the picture resulting from the sensitivity of the sensor (e.g. *minimum resolvable temperature difference* for infrared or contrast in brightness and hue for electro optical sensors). However, in this case it is of paramount importance that the decomposition of the criteria is complete. One arising problem at this level is the corresponding target function, which could not be determined in any case. For example, the image quality is among others dependent on the frequency inside the image. High frequencies imply an amount of features inside the image which could lead to false detections (*false positive*). Otherwise low frequencies could lead to an arising target leakage rate (*false negative*). Hence, the target function is neither maximization nor minimization, satisfaction or fixation. In case of calculating with an MSE, the frequency would be mapped in correlation with other up- and downstream modules (sensor, algorithm) to determine the intensity of the respective e*lementary criteria* (false detection or target leakage) at the *Measure of Effectiveness/Performance* level.

The *Measure of Effort* is used to determine the *System Effectiveness* in relation to the corresponding effort (e.g. costs, power consumption and required human interaction) which leads to the definition of *System Efficiency* as a utility/effort ratio and is considered at the *Measure of Efficiency*. Thereby the *Measure of Effort* is mapped to the *System Effectiveness* inside the criteria block (e.g. target nomination rate) to determine the *System Performance Potential* out of *System Effectiveness* and *Efficiency*.

4.2. The House of Quality

The traditional HoQ "*provides a method for a structured transfer from the costumers general, sometimes not well-defined, expressed needs and wants to the technical specifics*" [12]. To apply this approach to the task for evaluation of aerial platforms with regard to sensor surveillance capabilities the HoQ could be adapted in a way that it provides a method for the structured linkage from the general evaluation task (S&R mission) to related metrics. These linkages between the blocks as well as between the levels inside the blocks can be represented by dependencies at the room of the HoQ that is graphical and hierarchically structured and therefore easy to understand. The decision maker and all participating SMEs can walk through the house and identify as well as explain the background for each correlation or dependency and the linkages in between. This leads to a much more objective, transparent and reproducible decision result as well as to a significant simplification of complexity.

As multiple and interdependent criteria are needed for decision making, the commitment of multidisciplinary SMEs to the *Metric Derivation Process* is inevitable. Therefore a manageable design has to be provided which all participants can easily understand and work with. The basic structure of the HoQ leverages the integration of information from multiple sources and essential input parameters (e.g. target and utility functions) for multidisciplinary decision making.

4.2.1. Adapting the method

As shown in the previous chapter the HoQ at his initial design provides various attributes which can be transferred to system evaluation and analysis tasks. Nevertheless essential adjustments have to be conducted to establish an appropriate methodology which meets the complexity. The corresponding adjustments will be detailed below and are represented at Figure 4.

1) Defining the weights and the impact inside the room.

To fill in the room of the traditional HoQ, the decision maker tries to answer the question of, how much impact the design goals have in meeting the customer needs. As the final decision result should be able to be displayed on every level of detail at the block structure to increase transparency, the corresponding weights will change respectively. The traditional HoQ sums up the weights until the deepest level of decomposition is achieved. Therefore the first column listing the explicit scalar weights is obsolete. In the adapted structure the values inside the room does define the weighted composition of the parent criteria for the considered level of detail, whereby the values should be understood as related weights and not as dependencies.

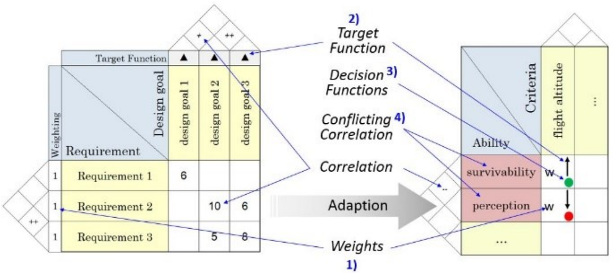

Figure 4: original HoQ (left) vs. adapted HoQ (right)

2) Consideration of multiple target functions for each parent-child combination.

The original HoQ provides a row between the roof and the room for itemizing the target function for each design goal answering the question whether a design goal has to be maximized or minimized to be optimal. At the *Metric Derivation Process* it is insufficient to provide only one target function for any node. With regard to that the target function can differ if one children has multiple parents with mutual target functions. For example, the capabilities of *survivability* and *environmental perception* inside the ability block are considered. Without representing the physical and mathematical equations it becomes obvious that both capabilities interact in an inversely way throughout the related distance. *Environmental perception* becomes optimal if the distance between sensor and object is minimal. Assuming that the object to be detected is a threat for the aerial platform (e.g. anti-aircraft artillery), the *survivability* simplify decreases by minimizing the distance. At the original HoQ the distance could have one target function and therefore the decision maker has to exclusive prioritize either *survivability* or *environmental perception*. To consider both in a compromising way, the structure is modified to provide a mapping of the target function separately for each parent-child combination inside the cells at the room (see Figure 4, arrows up = max and down = min).

3) Possibility to define and deposit relevant input parameters for decision theory.

Methods from decision theory are to be used in *Task Criteria Evaluation Process* (see Figure 1) and compute the provided output of the *Metric Derivation Process*. The traditional HoQ does not provide the possibility to define relevant decision functions which were subsequently used inside the *Task Criteria Evaluation Process*. Therefore the basic structure was modified to define all needed decision functions besides the weight and the target function inside the room. These functions depend on the different methods of *multi attributive decision theory* which can be, for example, utility functions, preference functions or cost functions (see dots at Figure 4, green: completely defined and red: not completely defined).

4) Possibility to consider correlations for elementary criteria quantification

The correlations inside the roof or the dooryard of the traditional HoQ can be used to identify critical design tradeoffs and comprised solutions. However, when doing so, the correlated scores are not used for further calculations. At the modified HoQ the identified correlations between the affected nodes are used to compile demands concerning the quantification of *elementary criteria*. Considering the capability example of *environmental perception* and *survivability*, it is not expedient to treat both capabilities independently and separately. On the contrary, it would lead to incorrect quantifications and possibly to irreversible decisions with far-reaching consequences. The aerial platform is located at a certain altitude at one time step and therefore both capabilities are directly connected through the distance. An implemented MSE for this interrelationship is presented in chapter 5.4.

As far as a correlation is defined and the combination of target function and correlation leads to a conflicting dependency, the quantification of both branches have to be obtained by comprehensive mathematical descriptions or simulations. In case no conflicting correlation exists, the quantification can be obtained by standalone criteria-oriented equations and simulations or by SMEs, if no mathematical solution is available. Nonetheless it should be emphasized the explicit advice to quantify the *elementary criteria* as objective as possible which means to use mathematical and physical background in terms of calculations and simulations.

The considered correlations provide also a second benefit. They could be used to identify potential for modification on the various alternatives which were evaluated and compared. For this the defined correlations can indicate critical criteria combinations or critical branches which could be optimized with minimal costs so that the decision maker, for example, can decide whether to invest money in modified mission sensors improving *perceptual capability* or even in noise reduction to improve *survivability*. The possibility to generate correlations for optimization (e.g. *Pareto Fronts*) out of the evaluation results still has to be reviewed and is not part of this work.

4.2.2. House of Metrics

Out of the various described modifications (Figure 4) on the original HoQ and to avoid misunderstandings concerning the range of application, the last term of the original method “*Quality*” is now renamed into “*Metrics*” which suggests the usage of this method for system evaluation instead of system design. The functional model of the *House of Metrics* (HoM) is shown in Figure 5. The model illustrates the affiliation of the different modules (top down block structure and described modifications from chapter 4.2.1) to the HoM through the background graphic. The input parameters are generated by the upstream *Stakeholder Expectations Process* in terms of evaluation task requirements. It becomes apparent that the metric structure consists of the criteria block structure and the related target function as well as the decision function which can also be seen as distance dimension. This should be understood beneath the term *metric* for the following executions. Furthermore it can be determined that the quantification function in combination with the resulting intensities of the *elementary criteria* are not part of the *House of Metrics*. Only input parameters for MSEs are provided in terms of e.g. *Measure of Effectiveness/ Efficiency/ Performance* without the corresponding value. The HoM provides further an approach to define correlated *elementary criteria* which leads in case of conflicting dependencies to the crucial use of interdisciplinary MSEs. Otherwise criteria-oriented MSEs can be used to quantify non-correlated *elementary criteria*. The substitutable calculation module to derive the intensity (MSE) is linked with the *elementary criteria* and could be replaced whenever a more suitable module is available. Summing up the HoM provides almost all relevant parameters needed at the further process (*Task Criteria Evaluation Process*) except of the criteria´s intensities which were calculated by substitutable MSE or – in case of necessity – determined by SME statements.

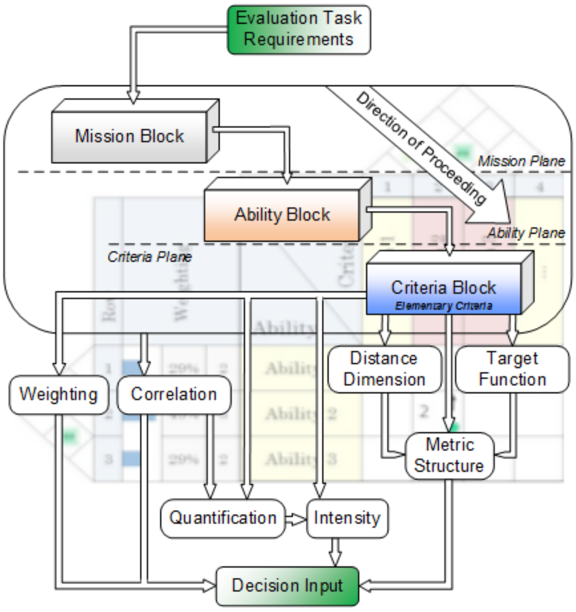

Figure 5: Functional model for the *House of Metrics*

5. METHODICAL IMPLEMENTATION

The considered capabilities in combination with the criteria decisive influence the result of the evaluation process. Thereby each elementary criterion has to be measurable. To demonstrate the general approach, an S&R mission is considered at the mission block. For general application it is of course possible to take several weighted missions into account.

To find the related nodes on the considered S&R mission a software-aided knowledge acquisition process was conducted with military and civil experts to identify the capabilities and their weighting structure on two levels inside the ability block for the related toy scenario.

5.1. Toy Scenario

The simplified scenario contains an aerial platform equipped with miscellaneous mission sensors which differ in their sensor parameters and connected image processing algorithms. As a target object a nonmoving person operating an anti-aircraft artillery (AAA) is assumed. The mission requirements are to conduct a long-term daytime observation by clear sky. Further the aerial platform has to avoid auditory detection and damage through the target object.

5.2. Criteria Catalogue for Environmental Perception Capability

The literature of target acquisition is often misleading or inconsistent due to the lack of a standard vocabulary leading to misunderstandings and ambiguous definitions of evaluation criteria. Therefore a glossary of search and target acquisition nomenclature has been developed in [13] to create a uniform foundation of definitions which were the most appropriate for description. These definitions were used for the corresponding nodes, while remaining or misunderstanding definitions have been determined with professionals out of the corresponding domain. To define the relevant nodes a software-aided application for knowledge acquisition was implemented for the usage by external SMEs. It should be highlighted that the participating experts are composed of civil and military full-time decision makers as well as tactical system analysts and had perennial experience in evaluation of aerial platforms. For deriving the fundamental capabilities to fulfill the defined scenario the experts worked separately with the provided application. They had the opportunity to select subjective important nodes from a predefined structure and the possibility to define new nodes in case of lacking relevant capabilities. Furthermore all selected nodes were well-defined or redefined to generate a consistent nomenclature with unambiguous definitions. An excerpt of the newly organized branch for the capability of *environmental perception* is shown in Figure 6. Due to the complexity of the shown excerpt the definition of the nodes inside the structure are not be listed.

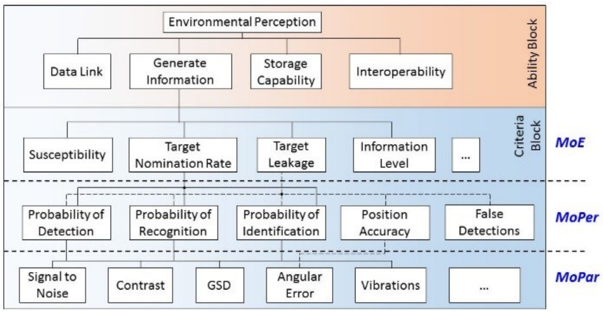

Figure 6: Excerpt of the organized sub-capability and criteria structure for the capability of *environmental perception*

As a child node from *environmental perception* one leaf is the capability to *generate information*. This capability combines much more sub-capabilities that could be defined in a third level inside the ability block or as metrics inside the criteria block. The work of the SMEs resulted in the decomposition into nodes like target nomination rate, target leakage and the provided *level of information*. These decomposed nodes have the sufficient level of detail to be quantified with mathematical equations and physical dependencies. So they become the first level of the criteria block (*Measure of Effectiveness*) which does not exclude the further decomposition. For example the target leakage is a criteria which could be simulated on a scenario related simulation if available or which could be decomposed further into the position accuracy, false detections and the probability of detection, recognition and identification (*DRI*) in combination with algorithmic performance. To illustrate the approach of the method (HoM) – especially the quantification of the *elementary criteria* – the probability of *DRI* and the provided *information level* will be considered in more detail in the following chapters, whereas further *elementary criteria* on other branches like endurance, data link or operability are excluded in the interest of clarity.

Just to demonstrate the general approach of the *House of Metrics* it is sufficient to map the quantification of the *elementary criteria* in a simplified way. Doing so, the performance of the processing algorithms is described in a simplified manner in this work, because the exponential increase in the number of sensors and information sources as well as perception algorithms requires some level of automation onboard the aerial platform which can only be represented in complex simulation models. One research focus at the Institute of Flight Systems (IFS) is "*Sensor- and Perception Management*" to select appropriate perceptional chains onboard the aerial platform to better accomplish ISR missions. To use the best chain, all eligible combinations are evaluated and compared with metrics. The applied image processing algorithms are dependent on several parameters like the used sensors and generated image quality (e.g. contrast, blur) as well as the processing power, weather conditions and the viewing angle which defines the shape of the object. However, the relation between these parameters and the quality of the processing algorithms is not fully understood and a current research topic at the IFS. It has to be stated, that simplified representations of complex dependencies is sufficient until more sophisticated MSEs are available. In general, the MSEs could be understood as interchangeable modules which can be replaced by much more high sophisticated calculation routines if available. Hence, for the simplified consideration of algorithmic performance in this work the contrast between the ground object and the surface background is calculated by brightness as well as hue and subsequently integrated at the simplified consideration of the *DRI probability*.

Despite the required objectivity for criteria quantification, in special cases the *elementary criteria* has to be subjectively determined like at the criteria for the provided *level of information*. This criterion evaluates the data processing output of a sensor chain by introducing a hierarchical scale depending on the provided *level of information* which is listed in Table 1. It is of perceptible difference for military decision makers whether only raw data (e.g. video stream or images) or symbolic information (e.g. regions of interest, highlighted objects or even situation assessment) are provided. Hereinafter the term "data" is used to a set of discrete, objective facts (e.g. image) and information is defined as categorized or contextual data (e.g. highlighted region of interest at an image or video).

lvl	Name	Description
0	Raw data	Provides raw image or video stream
1	Feature Assessment	Estimation of feature states and feature extraction (e.g. edge detection)
2	Area Assessment	Highlight featured areas without significant position accuracy (e.g. region of interest)
3	Detection Assessment	Estimate the feature as of mission interest using indirect, non-attributive information
4	Recognition Assessment	Specify the object through class discrimination (e.g. object is a human)
5	Identification Assessment	Specify the object through identification discrimination (e.g. sought person)
6	Tracking Assessment	Estimating trajectory of the object
7	Behavioral Assessment	Derive behavioral estimations through object motion
8	Intention Assessment	Derive inferences from estimated object behavior (e.g. person tries to hide)
9	Situational Assessment	Derive inferences from estimated attributes and relationships of entities (e.g. person is not threatening)

Table 1: Provided *information level* of data processing algorithms

5.3. Define the Weighting Vector

The literature to this subject provides a variety of different methods to determine weights in multicriteria models. According to the existing methods the most popular subjective methods are the direct ranking method, the pricing-out method and the *Analytic Hierarchy Process* (AHP). Unfortunately no statement can be made which method produces the most accurate results because “true” weights are and remain unknown. The detriments of most methods are firstly the necessary condition to rank the criteria on a ratio scale through setting weighting values at a defined interval which mostly leads to less confidence in the accuracy (statement from interviewed SMEs). Moreover, it was established that the human decision maker is more able to generate ordinal statements than to make pinpoint values. Secondly the ordinal rating with an equidistantly distribution (e.g. direct rating method, SWING method) is insufficient, because it is, however, necessary that the DM can define the distance between the criteria separately as not equidistant. Therefore an appropriate approach for defining the weights seems to be a combination of two methods, first to define the weights on an ordinal ranking scale through pairwise comparison with the AHP and subsequently adapting the distances in between by using the SIMOS method [14].

A direct ranking method as well as the combined method where implemented in a software-aided application and provided to the SMEs. At the direct ranking method the weights are determined through awarding priorities in terms of values on a scale of [0, 10]. This method is the most common applied in decision making processes and was used as reference. The test persons conducted the weighting process only for the first level of the ability block by using both methods. Subsequently both methods were evaluated. A little excerpt of the evaluation results are shown in Table 2.

	I have big confidence in the weighting results absolute true - not true				
M1		●●	●	●	
M2	●●●	●			
	The weights are transparent absolute true - not true				
M1		●●		●●	
M2	●●	●●			
	I would prefer M2 absolute true - not true				
	●●●	●			

Table 2: Results of weighting method evaluation: M1 = reference, M2 = combined approach

The SMEs quoted that the confidence in the established combined weighting system (M2) as well as the preference is much higher as on the reference (M1). According to the given answers and the evaluation of the combined method it can be determined, that the usage of the presented method to define the weighting system is a promising approach. It has to be emphasized that for this experiment the number of weighted criteria was only 6 and therefore manageable. By weighting the overall structure on both blocks and different levels the advantage of the combined method will be emphasized even more.

The SMEs defined the weighing vector for the situated criteria structure with the combined method by using a software-aided application. For clarity only an excerpt of the weighting results are shown in Figure 7. The left subplot shows the relevant capabilities for the toy scenario on the first level at the ability block. Meanwhile the right subplot displays the decomposed sub-capabilities (children) to the capability of *environmental perception* (parent). As every weighting system is dependent on the respective SME, certain deviations can be noticed. After a successful consistency check the distribution was averaged to define an unambiguous weighting system for the considered scenario and as a-priori knowledge. The resulting weights to the corresponding capabilities are shown in red.

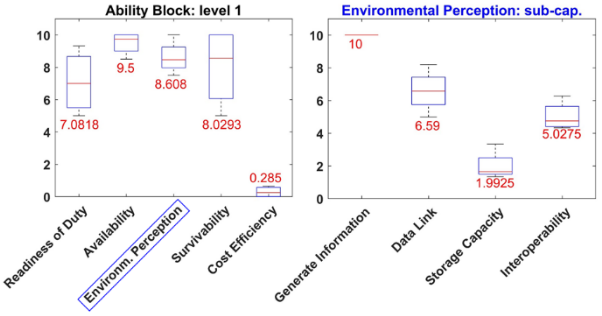

Figure 7: Excerpt of the defined weights from SMEs

5.4. Modelling and Simulation Environment

Figure 7 illustrates that the number of considered capabilities at the ability block in combination with the variety of the corresponding *elementary criteria* at the criteria block grows rapidly. Actually the quantification of these *elementary criteria* are found by subjective value allocation (compromised solution) or, as the preferred approach, by objective MSEs. However, if a conflicting correlation exists, a subjective value allocation is excluded in any case.

Hereinafter two significant interdepending capabilities are considered. As already pointed out in chapter 4.2.1 the *survivability* and the *perceptual capability* cannot be evaluated independently. For clarity and to show the approach of conflicting metrics only the capability to *generate information* as well as the *survivability* with the corresponding branches are examined in more detail below. Thus to calculate the intensity of the *elementary criteria* a simulation environment was implemented which contains a simplified point mass model of an aerial platform as well as a sensor and survivability model. The latter two are subsequently described. It should be pointed out that these following addressed models with their physical relationships and mathematical equations are greatly simplified.

Sensor performance model:

Literature provides a variety of different sensor performance models like the *NVL model*, *FLIR92*, *Thermal Range Model* (TRM4.1) and *NVThermIP* with different range prediction methods like the *Johnson* methodology and the *Task Targeting Performance* (TTP) metric.

Mostly the Johnson criteria is used to predict range performance and probability of target discrimination. Johnson published a table of needed cycles on target to conduct detection, recognition and identification tasks (*Measure of Performance*). Although more levels of discrimination for the Johnson criteria were defined, this methodology suffers from flaws [15]. Therefore the TTP metric is used in this work and implemented to predict the sensor performance characteristics. This metric explicitly takes into account the probability of detection independent of chance and dependent on the background contrast. The theoretical background of TTP is stated in [15]. TTP is calculated for the object shape in each dimension at the image plane (x,y) through coordinate transformation. According to the TTP approach task difficulty depends on how similar targets appear or the amount of target-like clutter in the background [16]. Therefore the contrast between the object (black) and the surface (yellow) is calculated by brightness (=226) and hue (=510) without considering atmospheric transmissions. For more precise calculations of the contrast the atmospheric transmission has to be considered using atmospheric models like for example *MODTRAN*. *Elementary criteria* on the *Measure of Parameters* level like the *signal to noise ratio* or the *minimum resolvable temperature* (MRT) could be calculated with the *Thermal Range Model* (TRM4.1).

As already mentioned current missions for time-sensitive ground object detection and tracking presuppose automated target recognition, guidance for autonomous vehicles, remote sensing and automated threat recognition. Therefore relevant data from different sources have to be identified and fused into a useful intelligence product [17]. The approach of the *Sensor- and Perception Management* presumes the consideration of all functional components like the sensor elements, the preprocessing algorithms and the corresponding resource management. Due to the fact, that no single algorithm is ideal under all circumstances, the sensor and the algorithm could not be considered independently. For that reason, although the simplified sensor performance model is sufficient for demonstration, it should be replaced for real decision tasks by holistic models which provide results in terms of the mentioned *elementary criteria* on the *Measure of Effectiveness/Efficiency* level.

Survivability model:

The survivability model contains three sub models for *detection avoidance* (e.g. electronic warfare, signature), *susceptibility* (e.g. shape, maneuverability, defensive aids) and *vulnerability* (e.g. critical component analysis, passive damage suppression). The *vulnerability* refers to the inability of the aerial platform to withstand the damage caused by the man-made hostile environment. The *vulnerability* is simplified represented by the shape of the aerial platform from the threats viewpoint. For concrete applications a complex critical component analysis as well as an endgame simulation is required. The detection module can be represented in different ways like as a radar station or even as human observation. At the related toy scenario the probability for *detection avoidance* is considered auditory through the emitted engine noise of the aerial platform and the surface arriving noise. In this respect a critical threshold for human hearing is defined with 54 dB which is obviously dependent on ambient noises of the environment. The probability of *auditory detection* is calculated through consideration of the systems engine noise and the atmospheric attenuation due to distance. The *susceptibility* is calculated for contact warheads like projectiles. The calculation rules are obtained from [18]. To calculate the number of projectiles striking the aerial platform, a miss distance model is used which is dependent upon the threat systems ability to track and guide the aircraft. Under the assumption of the considered toy scenario the threat is embodied by a single object (AAA). To simplify the calculation no ballistic flight behavior is respected. For each shooting attempt a Monte Carlo Simulation with an angular error of ±0.2° is conducted to calculate the probability a propagator hits the aerial platform. Therefore the real shape of the aircraft and the resulted hit probability are approximated by using a Carlton hit function [18].

For the toy scenario two systems are considered with different electro-optical mission sensors. An excerpt of the system (aerial platforms) parameters is listed in Table 3.

	System 1	**System 2**
Sensor EO		
Focal length (spotter)	600 mm	850 mm
Resolution	752 x 582	1920 x 1080
Sensor Size	36 x 24	36 x 24
Information output	Lvl. 0	Lvl. 2
Platform		
Size	7.2 x 6.4 x 2	14.8 x 8.2 x 2
Avg. engine noise @Loiter	110 dB	130 dB
Frequency @Loiter	150 Hz	150 Hz

Table 3: System parameters of both aerial platforms

The *elementary criteria* are calculated while the aerial platform is loitering stationary and horizontal above the target object with a constant turn radius (see Figure 8, platform = blue dot, target object = red dot). The dimension of the objects were approximated by length, width and height which leads to symmetric results for the two half circles. Therefore only the period for one half circle turn around the target object is considered (blue dotted line). The line of sight from the gimbaled sensor is always angled towards the object (yellow cone). The orange circle visualizes the area were noise detection is possible. The red circle represents the weapon engagement zone of the threat object (assumed 4000m). The right subplot represents the results of the Monte Carlo Simulation for 1000 shots on the platform. The green bars are the miss distances for the horizontal dimension and the dashed line represents the normative distribution for the miss distance. The blue highlighted area shows the dimension of the platform in horizontal direction. The miss distance is calculated with respect to the objects center. Thus it can be noted that even a miss distance of about 3m entail a hit on the platform. The miss distance for the vertical dimension is calculated equivalent.

The *DRI probability* for the first sensor system on the given flight path (blue line) is shown at the lower left subplot.

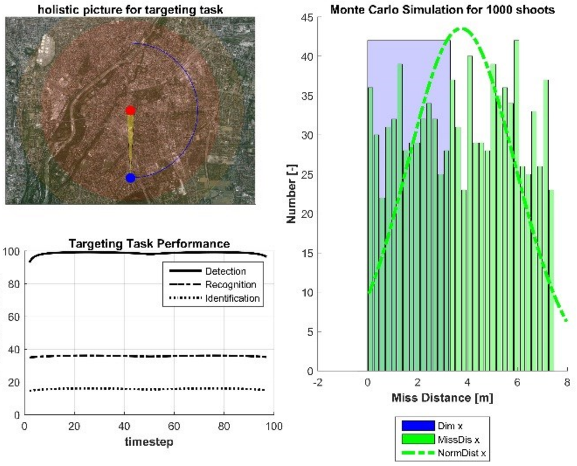

Figure 8: TTP metric for system 1 on semicircle flight path (blue dotted line) around the object with constant distance

Considering no static setup but periods of time the metrics could either be represented by a value over the time period (e.g. average or integral) or by the time the system fulfills the task at least with a satisfying probability or quality. To quantify the *elementary criteria* of the *perception capability* the integral is used to determine the normalized area under the curve of the TTP. The optimal intensity of TTP is 100 percent at every subtask (*DRI*). The referenced values are calculated from Equation (2).

$$I\begin{pmatrix} D \\ R \\ I \end{pmatrix} = \frac{\int_{t_1}^{t_2} TTP\begin{pmatrix} D \\ R \\ I \end{pmatrix} dt}{\int_{t_1}^{t_2} \begin{pmatrix} D \\ R \\ I \end{pmatrix}_{OPT} dt} \tag{2}$$

Due to the fact that optimal intensity for the *elementary criteria* of *survivability* is zero (no susceptibility or hits), the abovementioned equation cannot be used (division by zero). Instead the average of the time dependent probability is used to represent the intensity of the *elementary criteria*. The arising intensities are listed in Table 4.

Elementary Criterion	Intensity Sys. 1/2	Target Function
Capability: *Generate Information*		
Prob. of Detection	98,69 / 100.0	Max (100)
Prob. of Recognition	35,83 / 45.42	Max (100)
Prob. of Identification	15,75 / 28.21	Max (100)
Information level	0 / 2	Satisf. (lvl 5)
Capability: *Survivability*		
Detection Avoidance	19,47 / 38,71	Min (0)
Susceptibility	6,65 / 18,73	Min (0)
Vulnerability	40,39 / 46,87	Min (0)

Table 4: Results for conflicting *elementary criteria* of the capability to *generate information* and *survivability*

The target function for the *elementary criteria* of *survivability* is always the minimization whereas for the *DRI* task it is to maximize the probability. The optimal solution for the provided *information level* in the scenario is satisfied at level 5 which means specify the object (person) through automatically identification discrimination (see Table 1). This does not mean that the task is only accomplished at level 5 or higher like a constraint. It merely mean that all higher levels gain no more benefit because the mission does not require.
The additional *elementary criteria* are not considered in detail. Here simplified mathematical equations or subjective determinations are applied which is sufficient to demonstrate the holistic approach. It should be emphasized, that the quantification was conducted for electro-optical mission sensors. In case of additional consideration of thermal mission sensors, the quantification process has to be conducted respectively and merged to the result.

6. RESULTS ON TOY APPLICATION

The *House of Metrics* in combination with the quantified *elementary criteria* are the input parameters for the *Task Criteria Evaluation Process* (see Figure 9) which uses methods of *Multicriteria Decision Analysis* (MCDA) to rank the regarded systems with respect to the related mission. A variety of different methods are listed in [19] which can be used for detailed appraisal for a variable number of considered system variants. The used method to rank the systems for the considered scenario is *Promethee* which is based on pairwise comparison between the systems in every criterion. These comparisons take into account preference functions that determine the decision maker´s degree of preference between the criterion intensity of each pair of systems. Figure 9 shows the relevant input parameters defined inside the HoM (e.g. target function, weighting vector and intensity of *elementary criteria*) for the *Task Criteria Evaluation Process* as well as the preference functions concerning the considered *elementary criteria* of the previous chapter. The difference of the criteria intensity is plotted on the abscissa during the preference value is displayed on the ordinate. For the *probability of DRI* and *detection avoidance, susceptibility* and *vulnerability* the preference function flow is linear rising or rather falling respectively. The function flow of the *information level* shows the dependency to the satisfying target value. The broadest benefit would arise if one system has level 0 and the other level 5 or higher. As the difference amounts 2 levels the preference value arises to 0.4. The preference function is dependent on the corresponding target function and therefore adaptable on each cell inside the room.

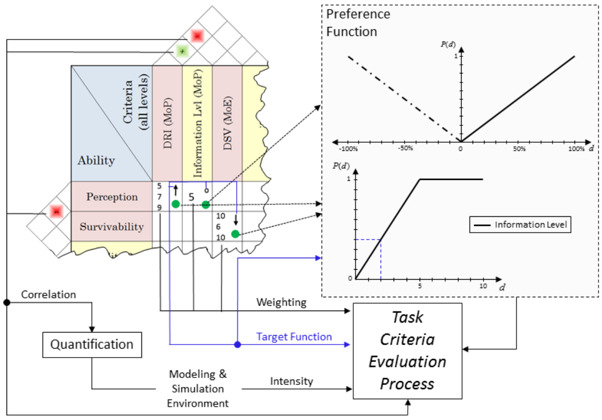

Figure 9: Input parameters for evaluation process provided by the HoM

For demonstration purposes the preference functions for the additional *elementary criteria* are assumed to be linear. The weightings for the considered *elementary criteria* are listed in the cells. It should be noted that the *probability of DRI* is a *Measure of Performance*, whereas the *probability of detection, susceptibility* and *vulnerability* is a *Measure of Effectiveness*. Both probabilities were actually listed in different HoM through they are located on different levels inside the criteria block.
The result of the *Promethee* method is presented in a degree of dominance whereas 100 percent means the system is dominant on every criterion (absolute dominance). A *System Performance Potential* for the first system of 46.85% and for the second system of 53.14% is calculated without considering any costs. Thereby it should be noted, that the more criteria are considered, the more the result will approximate to 50% if no system is absolutely dominant. An additional result is a prediction profiler of the *elementary criteria* as a function of weights shown in Figure 10. While the upper subplots show the sensitivity of the selected *elementary criteria* (*DRI* and *information level*) to the direct parent capability (*generate information)*, the bottom subplots indicate the sensitivity of these *elementary criteria* concerning the *System Performance Potential* to fulfill the S&R mission scenario. The plots show the

development of the results at different weightings [0,10] for the respective *elementary criteria*. The red dashed lines are the default weights. Thereby the weight changes only for the respective node during the remaining entries of the weighting vector stay constant. It can be noted that the weighting for *DRI* does not influence the ranking. Only the *elementary criteria* for the information level affects the *System Performance Potential*. If the level of provided information becomes not important (weight=0), the ranking will be inverted in favor of system 1 through the better survivability. Otherwise System 2 is always preferable through the better SPP. So the critical span of weights for a changing result were readily identified. With that the decision maker as well as the SMEs can examine carefully the robustness of the evaluation results. In case of sensitive weighting distributions through a ranking change the weights are of enormous importance and have to be checked carefully whether the assigned weighting value corresponds correctly to the related mission and are therefore appropriate to define the SPP.

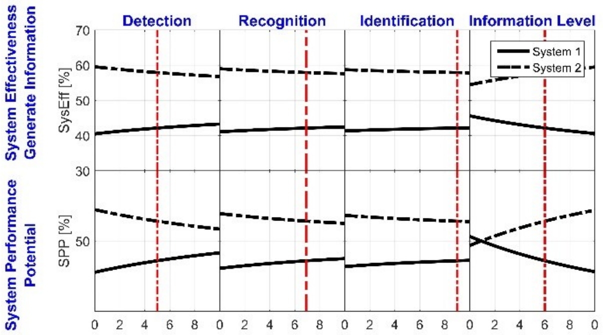

Figure 10: Sensitivity analysis for criteria weightings (x-axis) on the capability to *generate information* (upper plots) and the overall *System Performance Potential* (lower plots)

7. CONCLUSION AND FUTURE WORK

We presented an approach to derive a related metric structure in combination with relevant decision making parameters for a complex evaluation task. It was shown how the method (HoM) was derived and further applied on the toy scenario of an S&R mission. The method provides a solution for deriving mission-related metrics in combination with necessary input parameters for the subsequent evaluation process (e.g. utility-, target-, preference function) and demands for criteria quantification. Furthermore an excerpt of several branches of the metric structure for the capability of *environmental perception* as well as a corresponding MSE for the quantification of conflicting *elementary criteria* were presented.

An important next step – on which is being worked on currently – is to implement the overall concept for defining the *System Performance Potential* in an application to evaluate the holistic approach with civil and military SMEs. To evaluate the approach an evaluation process for an isolated mission will be considered, where external SMEs use the provided framework to conduct a complex system analysis task. The resulting proof of concept will be presented shortly.

Eventually investigations are planned on storing the SMEs knowledge using a fuzzy logic approach. In this work the criteria structure in combination with the corresponding weighting vector were extracted as values with certainty whereas the possibility to store the data as fuzzy values should also be taken into account and has to be evaluated.

8. REFERENCES

[1] F. W. Ploeger, “Strategic Concept of Employment for Unmanned Aircraft Systems in NATO,” Joint Air Power Competence Centre, Jan. 2010.

[2] R. Kumar, “Tactical Reconnaissance: UAVS versus Manned Aircraft,” Air Command and Staff College, AU/ACSC/0349/97-03, Mar. 1997.

[3] D. S. Soban, “A Methodology for the Probabilistic Assessment of System Effectiveness as applied to Aircraft Survivability and Susceptibility,” Georgia Institute of Technology.

[4] A. R. Habayeb, *Systems Effectiveness*. Burlington: Elsevier Science, 1987.

[5] R. J. Hillestad and L. Moore, *The Theater-Level Campaign Model: A Research Prototype for a New Generation of Combat Analysis Model*, 1996.

[6] F. A. Tillman, C. L. Hwang, “System Effectiveness Models: An Annotated Bibliography,” *IEEE Trans. Rel,* vol. R-29, no. 4, pp. 295–304, 1980.

[7] D. N. Mavris, Andrew P. Baker, and Daniel P. Schrage, “Implementation of a Technology Impact Forecast Technique on a Civil Tiltrotor,”

[8] D. N. Mavris, K. Griendling, and C. E. Dickerson, “Relational-Oriented Systems Engineering and Technology Tradeoff Analysis Framework,” *Journal of Aircraft*, vol. 50, no. 5, pp. 1564–1575, 2013.

[9] D. N. Mavris and K. Griendling, “Relational Oriented Systems Engineering and Technology Tradeoff Analysis (ROSETTA) Environment,” in *2011 6th International Conference on System of Systems Engineering (SoSE)*, pp. 49–54.

[10] S. Morawietz, M. Strohal, and P. Stütz, “Holistic Metrics for the Evaluation of Aerial Platforms,” DLRK 2015, Rostock, Sep. 2015.

[11] D. Baker *et al,* “Guidebook to Decision-Making Methods,” Department of Energy, WSRC-IM-2002-00002, Dec. 2001.

[12] J. B. ReVelle, J. W. Moran, and C. A. Cox, *The QFD Handbook*. New York: Wiley, 1998.

[13] D. Jones, “Air-To-Ground Target Acquisition Source Book: A Review of the Literature,” National Technical Information Service, Orlando, Sep. 1974.

[14] E. Siskos, “Elicitation of criteria importance weights through the Simos method: A robustness concern,” *European Journal of Operational Research*, vol. 246, no. 2, pp. 543–553, 2015.

[15] R. H. Vollmerhausen, “Target detection cycle criteria when using the targeting task performance metric,” in *European Symposium on Optics and Photonics for Defence and Security*: SPIE, 2004, p. 275.

[16] G. C. Holst, *Electro-optical imaging system performance,* 5th ed. Winter Park, FL, Bellingham, Wash.: JCD Publishing; SPIE, 2008.

[17] M. E. Liggins, D. L. Hall, and J. Llinas, *Handbook of Multisensor Data Fusion: Theory and Practice,* 2nd ed. Boca Raton, FL: CRC Press, 2009.

[18] R. E. Ball, *The Fundamentals of Aircraft Combat Survivability Analysis and Design,* 2nd ed. Reston, VA: American Institute of Aeronautics and Astronautics, ©2003.

[19] J. Figueira, S. Greco, and M. Ehrgott, *Multiple criteria decision analysis: State of the art surveys*. New York: Springer, 2005.