Simulation Based Performance Assessment - Methodology and case study

Dr Klaas Jan de Kraker, Wouter Noordkamp, Hilvert Fitski, Dr Ana Barros
TNO Defence, Security and Safety
The Hague, The Netherlands
klaas_jan.dekraker@tno.nl, wouter.noordkamp@tno.nl, hilvert.fitski@tno.nl, ana.barros@tno.nl

ABSTRACT

During the various phases of the Defense acquisition process, and in the early design phases, many decisions must be made concerning the performance and cost of the new equipment. Often many of these decisions are made while having only a limited view of their consequences or based on subjective information. Moreover, it is known that the impact of the decisions taken in the early design phases is large; they generally determine as much as 80% of the total life cycle costs. This highlights the need for decision making support in these areas.
To support decision-makers, during the various phases of the Defense acquisition process we introduce the Simulation Based Performance Assessment (SBPA) methodology. This methodology allows a transparent, unbiased and integral performance assessment of (future) platforms. It is based on Multi Criteria Analysis (MCA) and simulation techniques, and it considers the operational effectiveness, survivability, sustainability, and life cycle costs in the assessment. The methodology can be applied during acquisition of new platforms and systems as well as during maintenance and upgrade programs.

The developed SBPA methodology aims at integrally testing one or more system designs. The SBPA methodology supports:

1. integrally judging a single platform's design on its performance and comparing this performance with the life cycle costs;
2. comparing multiple platform designs on performance and life cycle costs.

This paper describes the SBPA methodology and the philosophy behind it. To illustrate its use, it also presents a case study that analyses and assesses alternative designs of a possible future platform. The case study involves the simulation of the platform's tasks from the perspective of operate, survive and sustain, and involves the calculation of its life cycle costs. This case shows that the SBPA methodology can be applied effectively to support making well-informed decisions during acquisition programs.

ABOUT THE AUTHORS

Dr Klaas Jan de Kraker is a member of the scientific staff at TNO Defence, Security and Safety. He holds a Ph.D. in Computer Science from Delft University of Technology. He has a background in Computer-Aided Design and Manufacturing, collaboration applications, software engineering (methodologies), meta-modeling and data modeling. Currently he is leading various simulation projects in the areas of simulation based performance assessment, collective mission simulation, multifunctional simulation and serious gaming.

Wouter Noordkamp graduated in Applied Mathematics at the University of Twente, Enschede, The Netherlands in 1998. Since then he has been working at TNO Defence, Security and Safety as scientist on the field of Maritime Operational Analysis. Wouter Noordkamp started working in the field of Above Water Warfare (AWW). He is currently working on Under Water Warfare (UWW) topics and the development of methods to assist the design process of new naval platforms. Furthermore Wouter is involved in studies in the field of reliability, availability and maintainability and the influence on the logistics. Wouter's specialties are statistics and signatures.

Hilvert Fitski studied Applied Mathematics at the Twente University of Technology in the Netherlands. After graduating and doing his military service, he worked at Philips Information Systems for two years. Since 1991 he has been working at TNO Defence, Security and Safety in The Hague.

Hilvert Fitski has been active in various areas supporting the Royal Netherlands Navy. First in the area of Above Water Warfare (AWW), mainly Anti Air Warfare (AAW), developing and using simulation models for the air defense of naval ships. Next his activities switched more towards Anti Surface Warfare (ASuW), mainly studies on surface surveillance for several types of crisis management operations and operations other than war. Further he is involved in Under Water Warfare (UWW), both Anti Submarine Warfare (ASW) and Submarine Warfare (S/M Warfare). In the latter field he coordinated studies for the four Walrus class submarines of the Royal Netherlands Navy.

At the moment Hilvert Fitski is program manager of the defense R&D programs "Simulation Based Performance Assessment (SBPA)" and "Underwater Defence in Expeditionary Scenarios".

Dr Ana Barros is Senior Research Scientist Military Operational Research at TNO Defence, Security and Safety and Senior Research Fellow at the Netherlands Defense Academy. She holds a Ph.D. from the Erasmus University of Rotterdam.

She has nearly twenty years of experience in the operations research field, and has been involved in a wide scope of military operation analysis projects varying from air defense, personnel planning, operations other than war, ground combat modeling, intelligence, logistics modeling, including strategic/operational ammunition requirements.

Simulation Based Performance Assessment - Methodology and case study

Dr Klaas Jan de Kraker, Wouter Noordkamp, Hilvert Fitski, Dr Ana Barros
TNO Defence, Security and Safety
The Hague, The Netherlands
klaas_jan.dekraker@tno.nl, wouter.noordkamp@tno.nl, hilvert.fitski@tno.nl, ana.barros@tno.nl

INTRODUCTION

The Dutch Ministry of Defence (NLD MoD) strives for an optimal balance between operational effectiveness and life cycle costs of its materiel. Because of the sea platform requirements of the Royal Netherlands Navy (RNLN), the RNLN fulfils the role of smart designer/smart integrator during the acquisition process. The RNLN acts as smart designer of its platforms and as smart integrator of the Sensor, Weapon and Command (SEWACO) systems.

During the acquisition process ('Defensie Materieel Proces', DMP), but also during maintenance/modernization programs, during the implementation of modifications and for doctrine-related questions, the NLD MoD uses simulation techniques. The use of these techniques is currently somewhat fragmented and is primarily aimed at the assessment of specific subsystems. However, simulation techniques are very suitable for integrated performance assessment, since they enable the evaluation of the system as a whole and reduce the risks and costs of real-life experimentation. The use of simulation methodologies for the integrated performance assessment improves the quality because the evaluation of the system as a whole, including its various subsystems, can be made beforehand.

In order to support better decision making during the acquisition process, we have developed the Simulation Based Performance Assessment (SBPA) methodology. The SBPA methodology enables the determination of the overall performance of a platform. It considers the operational effectiveness, survivability, sustainability, and life cycle costs. These aspects are compared in a transparent and unbiased manner. In this way, the SBPA methodology enables to perform trade-offs between different design alternatives and it allows making well-informed acquisition decisions.

This paper presents the SBPA methodology and illustrates its application using a case study. As such, we start by presenting the different phases of the SBPA methodology and steps within these phases. In order to illustrate the SBPA methodology, we present a case study that applies the methodology on a (fictitious) frigate case. This case study analyses the performance and costs of two alternative frigate designs. It illustrates the different steps and techniques of the methodology. In particular, it highlights certain areas of interest: the determination of performance and costs and comparison of alternatives.

Together the methodology and case study descriptions provide a global overview of the SBPA methodology and provide more detailed insight in certain techniques applied in the methodology.

It must be noted that the sole purpose of the case study is to demonstrate the methodology. The case study uses fictitious platform designs and parameters. It does not reflect decisions or intentions of the RNLN in any way. No conclusions can be drawn from the study results other than regarding the use of the methodology.

The remainder of this paper is organized as follows. The next two sections introduce the SBPA methodology and present the results of our frigate case study. Lastly, we present our conclusions and lessons learned.

SBPA METHODOLOGY

This section introduces the SBPA methodology and shows its relation to other approaches.

The SBPA methodology allows a transparent, unbiased and integral performance assessment of (future) platforms. It defines an evaluation framework to assess the overall performance of a platform to support the acquisition process. It is based on Multi Criteria Analysis (MCA) and simulation techniques.

The US DOD has developed the Cost and Operational Effectiveness Analysis (COEA) approach, see [1]. The goal of COEA is to perform a cost-benefit analysis to support acquisition decisions. Important aspects of the COEA approach are the problem definition, determining alternative system concepts, defining system requirements and MOEs, defining scenarios, data collection and model selection, and estimating costs and operational effectiveness of the possible system concepts.

Based on the COEA, the Integrated Cost and Operational Effectiveness Approach (ICEA) was developed, in the NATO Future Reduced Cost Combatant study, see [2]. The ICEA analyses design alternatives in the early design stage using operational effectiveness and life cycle cost. It defines scenarios and missions, and it develops a mission task tree. The operational effectiveness is calculated using the Operational Value Model (OVM), see [3]. The survivability and sustainability are calculated separately. ICEA combines the operational effectiveness and the survivability and sustainability into an overall Figure of Merit (FOM) for the platform using the MCA Analytic Hierarchy Process [4].

Both OVM and ICEA use an aggregation method to combine the different MOEs into operational effectiveness and an overall performance of the platform. These aggregation methods are common to Multi-Criteria Analysis (MCA) [5]. MCA is a scientific evaluation method for making rational choices between several discrete alternatives. It evaluates and compares the alternatives to obtain an aggregated assessment which results in a ranking of the alternatives.. Typical for most MCA processes are preparing an evaluation structure (model) and dealing with the subjective preferences of one or more stakeholders involved and the processing of information from diverse and diffuse nature and origin (conflicting, quantitative and qualitative, inexact and uncertain, etc.).

The SBPA methodology leverages these approaches and uses the following elements particularly to establish a structured process:

- the COEA approach
- mission-task analysis (ICEA and OVM)
- simulation models with different levels of aggregation and fidelity
- MCA techniques for weighing and comparing the different components

The SBPA methodology consists of five phases each consisting of several steps, see also Figure 1. Here a short overview of the five phases is given. A detailed description of the step by step plan can be found in [6].

Phase 1: Definition phase

In this phase the operating conditions of the platform are defined: the expected missions and tasks that will be carried out by the platform under consideration as well as the expected scenarios in which the platform operates are described. Since exercises are an important part of the usage of a platform, they should be taken into account to properly determine sustainability and cost

During the definition phase a structured description of the assumptions made and expected scenarios is produced.

Phase 2: Product effectiveness

In the second phase the evaluation framework is developed. An important part of the evaluation framework consists of determining the overall effectiveness of the platform, the so-called product effectiveness. Product effectiveness consists of three parts: operational effectiveness, survivability and sustainability. These are defined as depicted in Table 1.

Table 1. Product effectiveness parts.

Operational effectiveness	The ability to perform military, diplomatic, police and humanitarian missions and tasks
Survivability	The ability to sail and fight after suffering an attack and the accumulation of battle damage
Sustainability	The ability to ensure the availability of propulsion, maneuverability, weapon, sensor and support systems to keep the platform operational

For each of these parts, the criteria that are to be evaluated are placed in an MCA tree. An MCA tree contains criteria on which a design will be assessed and the relationships between the various criteria. For the operational effectiveness, this tree is usually called a mission task tree, because this tree describes the relations between scenarios, missions and tasks and their relative importance, given by weights that are determined by expert opinion. For sustainability, criteria such as reliability, maintainability, required personnel and spare parts are commonly used.

For each criterion, one or more Measures of Effectiveness (MOEs) are formulated to evaluate the performance of the platform on this criterion. Also, it is determined how the MOEs are translated to a common scale, by using utility functions. In this way, MOEs can be meaningfully combined.

This evaluation framework is setup in advance, in order to avoid bias in the evaluation process. This framework enables a smooth evaluation process for the various design variants. Because all evaluation aspects are now described, new design variants can be added later on, following the same procedure. Moreover, it enables a fair comparison of alternatives.

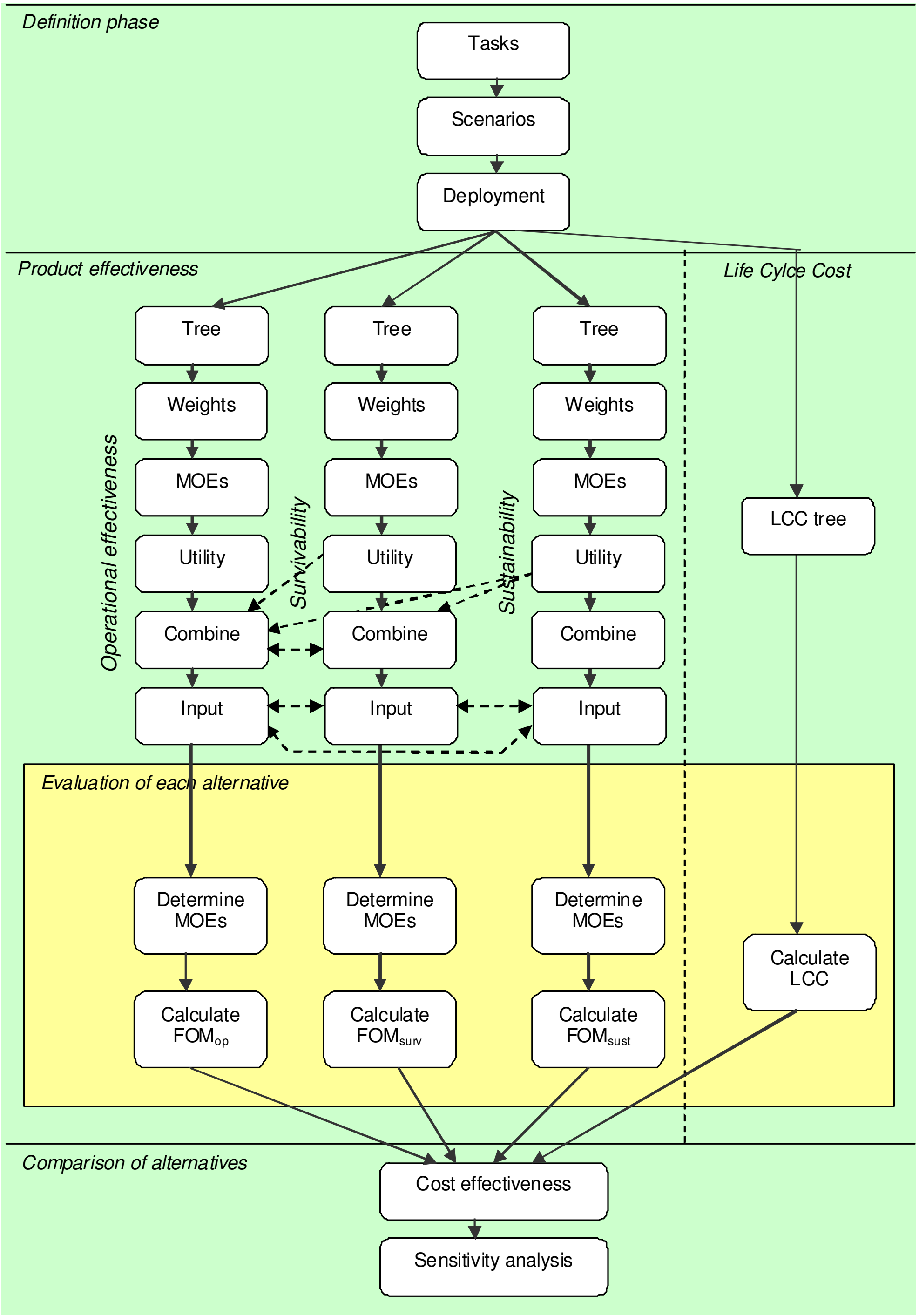

Figure 1. Step by step plan.

A part of the setup of the evaluation framework is the decision how the product effectiveness of a design is determined, i.e. using expert opinion or simulation. Simulation models are an invaluable tool. Careful consideration is needed regarding the choice of input parameters and variations in the simulation input to ensure that the simulation results are representative for the real effectiveness of the platform.

Phase 3: Life Cycle Costs

The third phase defines how the Life Cycle Costs (LCC) are determined [7]. The LCC of a new platform can be divided into 6 main categories:

- Design
- Development
- Production
- Operating
- Support
- Disposal

Each category consists of a number of subcategories for which the costs can be determined in various ways and with more or less detail.

The three cost estimation methods that are used mostly are Analogy, Parametric and Engineering.

The analogy method compares a new system (design) with one or more existing systems and does not only take into account technical differences, but also differences in use and/or maintenance concepts.

The parametric method estimates the cost based on various measurable properties of the system. It is based on the existence of a causal relationship between system cost and system properties.

Unlike the parametric method, in the engineering method the formulas are based on the detailed consideration of the system and its use, and the relations between them. This method requires detailed design information.

Phase 4: Evaluation of each alternative
The first three phases have defined the evaluation framework that will be used in this phase to assess the performance and life cycle cost for each of the alternatives. In this way the alternatives can each be assessed in an identical manner.
This phase determines the actual performance for each of the alternatives. Using the results of phase 2, it calculates values for each of the defined MOEs, and it aggregates these MOE values into a Figures of Merit (FOMs) for each of the three product effectiveness parts. This may involve simulation as well as using expert opinions. And, using the results of phase 3, it performs the calculation of the LCC.

Phase 5: Analysis of results and comparison of alternatives
The final phase is used to analyze the results, compare the alternatives on product effectiveness and costs and carry out a sensitivity analysis to determine the robustness of the results.

All steps are summarized in Figure 1.

CASE STUDY

The structure and steps of the SBPA methodology have been tested by applying the methodology to a case study. This case study also illustrates the SBPA methodology.

Variants of a fictitious new frigate have been designed that can be used for both Above Water Warfare and for Underwater Warfare purposes. These variants vary from a relatively simple design up to a more sophisticated frigate. Note that in a real design process, numerous design variants will be considered with smaller differences between them.

The more sophisticated frigate has better sensor and weapon systems, a lower visibility on radar and a higher maximum speed.

Phase 1: Definition phase

In this phase the future deployments of the ship are defined: its missions and tasks, the scenarios in which the ship is expected to operate and how these deployments are divided over the life time of the frigate.

Step 1: Missions and tasks
The missions of the ship are defined to be:

- Humanitarian aid and disaster relief
- Coast guard
- Sea control
- Escort of a High Value Unit
- Anti Submarine Warfare
- Mine counter measures

Each mission consists of one or more tasks.

Step 2: Scenarios
The ship is expected to operate in four different scenarios with varying level of violence:

- Humanitarian operation after a hurricane in the Caribbean
- Coast guard in the Caribbean
- Conflict situation in the Middle East.
- Conflict situation in the Baltic Sea.

Step 3: Life cycle description
The expected operational life time of the ship is 25 years. During this life time the ships operates in the defined scenarios and carries out the missions described. Large maintenance is carried out regularly. After this maintenance period of several months, a period of several months is reserved for training of the crew.

Phase 2: Product effectiveness

In the second phase, the framework for the evaluation of the product effectiveness is constructed. Product effectiveness is given in three separate measures for operational effectiveness, survivability and sustainability. In this paper, only the construction of the operational effectiveness part of the product effectiveness is illustrated. However, the survivability and sustainability can be determining analogously.

Step 4: Mission task tree

A mission task tree is constructed based on the results from the previous two steps, see Figure 2.

Step 5: Weights

Weights of the scenarios, missions and tasks in the mission task tree are determined by experts and are based on their relative importance.

Weights of tasks are determined per mission. While determining the weights, experts could take into account that tasks for which the platform is specifically designed get a higher weight. Also the expected number of times a ship will be carrying out a certain task could be taken into account.

The weights of the scenarios and the missions are determined in the same way.

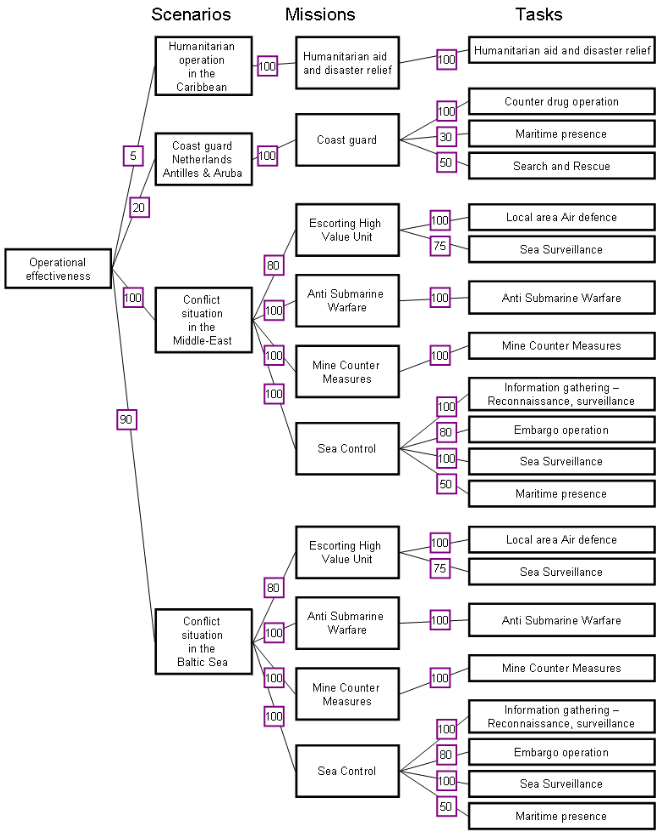

Figure 2. Mission task tree with weights.

An effective way to determine the weights is to assign 100 points to the most important task per mission and rate the other tasks relative to these 100 points. The weights are normalized by dividing the points per task by the total number of points of all tasks within a mission.

The resulting mission task tree with assigned weights is depicted in Figure 2. Here, for example, the scenario ‘Conflict situation in the Middle-East’ is assigned a high weight, because the platform has been designed for such conflict scenarios. The task ‘Maritime presence’ within the mission ‘Coast guard’ is assigned a relatively low weight, because it is more a supplementary task that can also be carried out by other platforms.

Step 6: MOEs

For each task, Measures Of Effectiveness should describe the performance of the platform. In some cases one MOE will be sufficient, but it is also possible to define more than one MOE and to combine them later on (during step 7).

The case study defined MOEs for each task, for example:

Humanitarian aid and disaster relief: the number of persons that can be evacuated, the time needed to arrive in the disaster area, the number of persons that survive the disaster thanks to the coordinated intervention.

Counter drug operation: the number of drug smuggling boats that have been prevented to deliver the drugs.

Search and Rescue: the time needed for a search and rescue (SAR) operation.

Local area Air defense: probability that a HVU is not hit by an enemy air attack.

Sea Surveillance: percentage detected and correctly identified ships.

Step 7: Utility functions

The MOEs mentioned under step 6 show that there is a large variety of possible MOEs. Combining these MOEs into a single value should be done in a sensible way. For example, it is a bad idea to simply add up the time needed for a SAR operation and the number of drug smuggling boats that have been caught. Utility functions can be used to make MOEs easily comparable and to combine them.

Utility functions are most known in the economic theory, but are also used in other scientific fields ([8], [9]). With a utility function, MOEs can be translated to easy understandable values, the so-called scores, that have the same meaning and the same scale for each criterion. This enables not only the meaningful combination of

MOEs, but it can also retain the meaning of a score at a combined level.

We propose utility functions that translate the MOE to a score between 0 and 10, with 10 the best score and 0 the worst.

Figure 3 shows the utility function used for the search time in a SAR operation. As fast rescue is essential for a successful SAR operation, a high score is assigned if the search time is short. However, if the search time increases, the score quickly decreases.

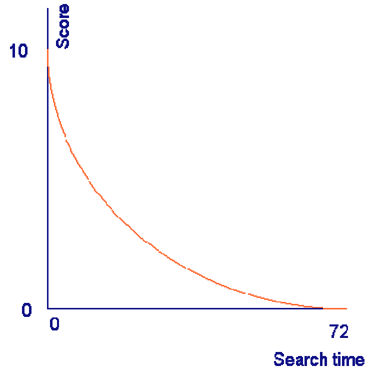

Figure 3. Example of utility function for SAR.

Another utility function has to be used to translate effectiveness in an escort operation. If the effectiveness represents the probability that a HVU is not hit by enemy fire, then the higher this probability, the better. It is reasonable to assume that a certain minimum effectiveness is required, otherwise the HVU will not be escorted by the protecting ship. Such assumptions lead to the utility function in Figure 4.

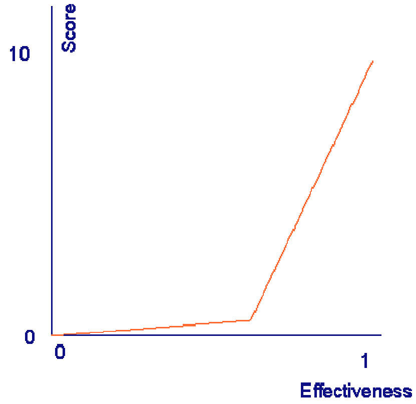

Figure 4. Example of utility function for Local Area Air Defense.

Step 8: Combination of MOEs

After MOEs have been translated into scores, these scores can be combined. We use the well-known MCA aggregation method, the weighted sum which is given by

$$
\begin{aligned}
FOM &= \sum_i \sum_j \sum_k s_i m_{i,j} w_{i,j,k} S_{i,j,k} \\
&= \sum_i \sum_j \sum_k s_i m_{i,j} w_{i,j,k} g_{i,j,k} (MOE_{i,j,k})
\end{aligned}
$$

With FOM the total Figure of Merit, $w_{i,j,k}$ the weight of task k in mission j and scenario i, $m_{i,j}$ the weight of mission j in scenario i and s_i the weight of scenario i in the FOM, $S_{i,j,k}$ the score of task k, $MOE_{i,j,k}$ the value of the MOE of task k and g_k the utility function to translate the MOE of task k, all in mission j and scenario i. If there are multiple MOEs per task, the equation can be adjusted accordingly.

A weighted sum can be used if all criteria are independent of each other, which means in our case study that the tasks are independent of each other. This is only true if tasks are not carried out simultaneously or one task needs to be completed successfully before another task can be carried out. The SBPA methodology gives suggestions how to deal with dependencies. However, in our case study, all criteria can be assumed to be independent.

A common problem of using the weighted sum as aggregation method is that it can yield to so-called compensation effect. A compensation effect implies that bad scores can be compensated by good scores on another task. Scores on individual tasks might get lost if the comparison between alternatives is done only on the Figures Of Merit. In many cases bad scores on a single task are worth to pay attention to. Therefore the SBPA method uses knock out criteria. With a knock out criterion, a minimum score on a task is required. If a design alternative does not meet this minimum score, it will be eliminated from the comparison of alternatives.

Step 9: Simulation input

First, it has to be decided in which way the MOEs will be determined, e.g. which simulation model is to be used.

The definition of the simulation input is very important, because the value of the input parameters has to be chosen such that they form a representative set of scenario variations. Variations that we took into account are: differences between summer and winter, different types of air threats, different intention of drugs smugglers et cetera.

Most simulation models will be using Monte Carlo simulation, i.e. multiple replications are carried out to account for stochastic elements in the model, such as detection probabilities, probabilities of weapon effectiveness and (random) behavior of the opponent.

It is important to carefully consider the required number of replications. Too little replications cause inaccurate results, but too many replications cost unnecessary time, therefore delaying the assessment process.

The simulation output should not only lead to a reliable estimation of the mean of the MOEs, but also to a reliable estimation of the distribution of the MOEs. This distribution will be used in the sensitivity analysis in phase 5. The number of replications should be large enough for both estimations.

Phase 3: Life Cycle Cost

Step 10: Life Cycle Cost

For the determination of the Life Cycle Cost, the analogy method has been used. The costs of existing Dutch frigates have been used for this analogy. These costs have been multiplied by a certain factor per cost element that takes inflation and costs into account as well the specific new characteristics of the new designs. If the costs of a certain category were unknown, they have been estimated based on common ratios between cost categories.

The life cycle costs of the designs per category were estimated as given in Table 2.

Table 2. Estimation of LCC.

Cost Category	**Design 1 [M€]**	**Design 2 [M€]**
Design	22	33
Development	33	50
Production	443	643
Operating	466	512
Support	144	144
Disposal	0	0
Total	**1108**	**1382**

The disposal costs are assumed to be 0. Although the ship may be sold instead of broken up, this will involve additional sales costs which reduces any possible profit.

The extra costs of design 2 are caused by the better design. Especially increased maximum speed will be a cost driver. The extra costs are reflected in the design and development phase, because the extra requirements and new systems lead to a larger design and development process. More advanced systems lead to higher production cost and higher operating costs (e.g. because of higher maintenance requirements).

Phase 4: Evaluation of each alternative

Step 11: Calculation of MOEs

If the set up in phase 2 and 3 has been carried out and described with enough detail, this set up can serve as a recipe for the determination of the MOEs per design alternative.

This means running the required simulation models with the given input variations and elicitating the required experts. The MOEs can be translated to scores with the help of the utility functions.

Step 12: Calculation of FOM

The MOEs are combined to obtain Figures of Merit (FOMs) that represent that product effectiveness.

We present the scores on operational effectiveness as a traffic light board in Table 3. This can be used for quick insight in the scores of the missions and tasks. Scores get a color red, yellow or green so that low or high scores can be signaled quickly. The scores corresponding to the colors are advised to be: 0-4 (red), 4-7 (yellow), 7-10 (green).

Table 3. Scores of tasks and traffic lights.

Item	Weight	Score design 1	Score design 2
FOM		6.9	7.4
Scenario 1: Humanitarian operation in the Caribbean	0.023	7.2	7.2
Mission: Humanitarian aid and disaster relief	1.000	7.2	7.2
Humanitarian aid and disaster relief	1.000	7.2	7.2
Scenario 2: Coast guard NA&A	0.093	6.8	7.7
Mission: Coast guard	1.000	6.8	7.7
Counter-drug operation	0.555	6.2	7.8
Maritime presence	0.167	7.0	7.0
Search and rescue	0.278	8.0	8.0
Scenario 3: Conflict situation in the Middle East	0.465	6.9	7.3
Mission: Escorting HVU	0.211	7.9	8.9
Local area air defence	0.571	8.1	9.4
Sea surveillance	0.429	7.7	8.2
Mission: ASW	0.263	4.0	4.8
ASW	1.000	4.0	4.8
Mission: Mine countermeasures	0.263	8.0	8.0
Mine countermeasures	1.000	8.0	8.0
Mission: Sea control	0.263	7.8	7.9
Information gathering	0.303	7.6	7.6
Embargo operation	0.242	8.5	8.6
Sea surveillance	0.303	7.7	8.2
Maritime presence	0.152	7.0	7.0
Scenario 4: Conflict situation in the Baltic Sea	0.419	7.0	7.5
Mission: Escorting HVU	0.211	7.9	8.9
Local area air defence	0.571	8.1	9.4
Sea surveillance	0.429	7.7	8.2
Mission: ASW	0.263	4.6	5.5
ASW	1.000	4.6	5.5
Mission: Mine countermeasures	0.263	8.0	8.0
Mine countermeasures	1.000	8.0	8.0
Mission: Sea control	0.263	7.8	7.9
Information gathering	0.303	7.6	7.6
Embargo operation	0.242	8.5	8.6
Sea surveillance	0.303	7.7	8.2
Maritime presence	0.152	7.0	7.0

Note that traffic lights, utility functions and knock out criteria can all be used to detect and process (intermediate) inferior results.

Phase 5: Comparison of alternatives

In our case study, the two design alternatives showed FOMs and costs as in Table 4.

Table 4. Comparison of alternatives.

	Design 1	Design 2
Operational effectiveness	6.9	7.4
Survivability	5.7	6.1
Sustainability	6.0	5.9
LCC	1108	1382

The operational effectiveness and survivability of design 2 are higher. However, the costs are higher as well and the sustainability is lower.

Various methods exist to determine the cost effectiveness of design variants, e.g. dividing the effectiveness by the costs. However such a method has a major drawback as it assumes linearity between effectiveness and cost. Usually this assumption does not hold. Sometimes a cheap but inferior product can have a favorable cost effectiveness ratio. On the other hand, a product with a favorable cost effectiveness ratio can exceed the available budget. Bounds on effectiveness and costs are therefore very important. Another disadvantage is that the operational effectiveness, survivability and sustainability need to be combined in a total product effectiveness figure, losing the insight in the separate aspects.

It is better to look at requirements to product effectiveness and costs, e.g. discarding alternatives with costs that exceed the maximum budget or an aspect of the product effectiveness that is below the minimum required effectiveness.

This will lead to a number of resulting alternatives. Assessment of these alternatives will not always be based just on a rational method, which makes it difficult to give a method that determines *the* best alternative.

Variants can be assessed on various aspects. Variants that are better or worse on all aspects can easily be identified. If variants are better on one aspect and worse on another, such as in this case study, it depends on the weight that is assigned to an aspect which variant is better. However, often these weights are not known in advance.

Therefore we propose the following method to assist in this assessment.

For an assessment on operational effectiveness and costs, define the overall score of a design by $wE + 10(1-w)(1-C/B_{\max})$, with w the – varying – weight of the operational effectiveness, E the operational effectiveness, C the cost and B the maximum budget, in this case study assumed to be 1500 M€. Varying the weight w between 0 and 1 leads to the assessment shown in Figure 6. On the left, the LCC are all-determining, on the right it is the performance that determines the score.

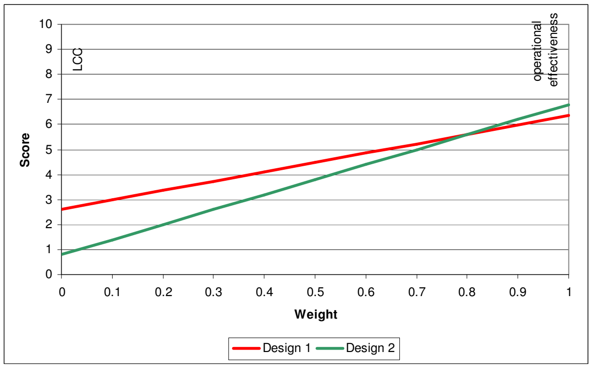

Figure 6. Costs versus operational effectiveness.

This picture shows that unless the weight of the effectiveness is relatively high, design 1 (the cheaper alternative), is preferred.

This method can be extended by comparing the other aspects (survivability and sustainability) pair wise as well.

Figure 7 shows the results.

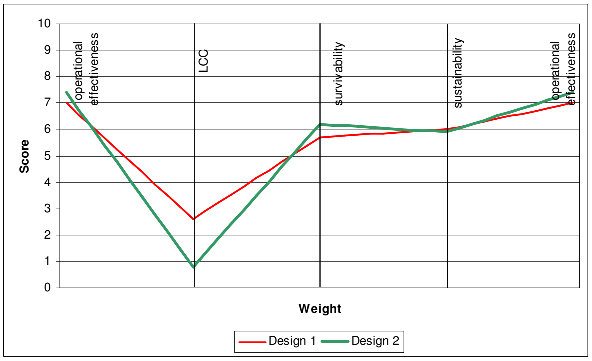

Figure 7. Comparison of alternatives.

<u>Sensitivity analysis</u>

An important step for the comparison of alternatives is a sensitivity and uncertainty analysis. When using stochastic simulation models one has to take into consideration the inherent statistical variation. Besides this source of uncertainty, one should also take into account the input data uncertainty and the uncertainty introduced by expert judgments.

A way of dealing with these sources of uncertainty is the use of Monte Carlo analysis. In such an analysis, the uncertainty of the scores is modeled using a statistical distribution.

In the Monte Carlo simulation, N replications are carried out. In each replication, a value is generated for each score according to the distribution of that score. The resulting values are aggregated as if they were the scores of the alternative designs, leading to an overall score on effectiveness and costs. In each replication, the scores and the ordering of alternatives are determined. The aggregate results over the N replications give information on the robustness of the scores and rankings.

In order to apply the method described above, the statistical distribution associated to the scores has to be established. Different approaches are used for this depending on whether simulation or expert judgment is applied.

For simulation results, information regarding the distribution of the output can be available as output of the simulation, e.g. an empirical distribution based upon the simulation outcomes (each replication is part of the distribution).

In the case that the scores are obtained via expert judgment, the following approach is used. Experts are asked to indicate their uncertainty regarding their opinion, e.g. although they expect a score to be 8, it likely that the score will be between 6.5 and 8.5. This uncertainty is modeled using the triangular distribution given by

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)} & a \le x \le c \\ \dfrac{2(b-x)}{(b-a)(b-c)} & c \le x \le b \\ 0 & otherwise \end{cases}$$

With c the expected score, a the lower bound and b the upper bound given by the experts.

In the case study 10,000 replications have been carried out, leading to the following distributions of the Figures Of Merit representing the operational effectiveness of design 1 (Figure 8) and design 2 (Figure 9). Mean, median and standard deviation of the FOM_{op} of both designs can be found in Table 5.

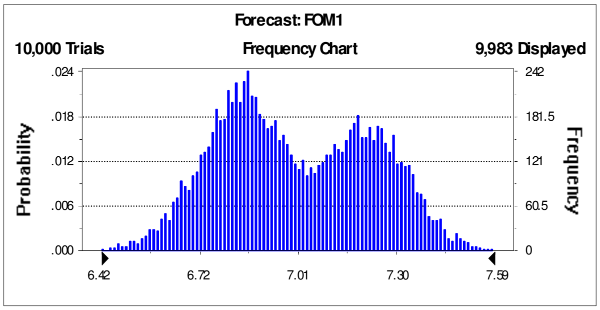

Figure 8. Distribution of FOM_{op} design 1.

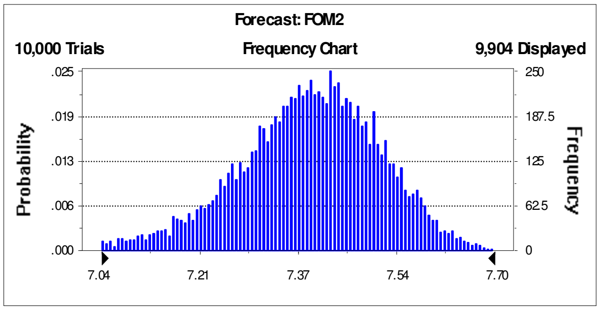

Figure 9. Distribution of FOM_{op} design 2.

Table 5. Results Monte Carlo simulation.

	Design 1	Design 2
Mean	7.00	7.39
Median	6.97	7.40
Standard deviation	0.23	0.12

For the other criteria (survivability, sustainability and costs), the same analysis can be carried out.

Ranking of alternatives

The alternatives can be ranked as well. Figure 10 shows the ranking of the two designs. The colors indicate the ranking order: the best alternative (green) and the second best alternative (yellow). Figure 10 shows that the probability that design 2 is the best is 88%.

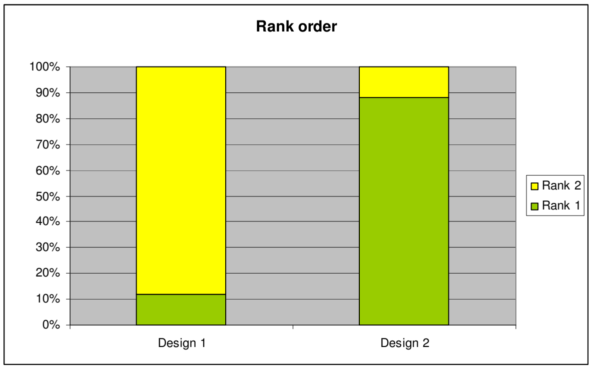

Figure 10. Ranking orders.

The Monte Carlo simulation should be set up in such a way that in each replication the same condition should apply for each of the designs. E.g. if the operational environment might vary with regard to the radar conditions - leading to a varying radar detection distance and therefore varying operational effectiveness - then the same conditions should apply per replication for both design 1 and 2.

Therefore simply drawing two random numbers from the defined distribution will not work properly, the random streams should be recreated for the applicable parameters or the designs should be simulated at the "same time".

CONCLUSIONS

This section presents our conclusions of using the SBPA methodology. It first presents our general conclusions followed by a summary of the main lessons learned.

We have presented the SBPA methodology and its application in a case study. The purpose of this was to describe the SBPA methodology and to illustrate its intended use in an acquisition process.

The SBPA methodology is a workable and practicable methodology to assist decision makers during the acquisition process. The SBPA methodology specifies a step-by-step plan that is hierarchically structured in phases and steps. It addresses operational, survivability, and sustainability aspects as well as life cycle costs. It enables the execution of a performance assessment of a platform design.

Although the SBPA methodology has clear advantages it does not provide an easy recipe that can be followed blindly. To aid future studies using the SBPA methodology we here summarize our main lessons-learned.

- Although it may be difficult, it is important to describe the operational context of the platform adequately. This includes for example, the life cycle, characteristics of the areas of operation and of the behavior (tactics, level of aggression) of opponents. Care must be taken that the assumptions that are made have a broad coverage and are plausible at the same time.
- As applying the SBPA methodology is quite time consuming, it is recommended to determine timely the required level of detail. For example, if the contribution of one task to the overall FOM is very low, it should be considered to address this task in less detail, i.e. with less effort.
- The Definition phase specifies the input for generating MOEs for the different tasks. However extra assumptions have to be made when setting up and executing the simulations, for example, many additional assumptions are made regarding environmental conditions and system parameters. It is recommended to record these assumptions and to assure consistency among them.
- We have experienced that the sensitivity and uncertainty analysis during the comparison of alternatives is very useful. It provides insight in the robustness of the performance assessment.

All in all, the SBPA methodology is a suitable methodology for executing performance assessment of new platforms and systems as well as during maintenance and upgrade programs. Based on our experiences we believe that the use of the SBPA methodology leads to better choices and improves the transparency during the acquisition process.

REFERENCES

[1] M.W. Garrambone (1994). An overview of Cost and Operational Effectiveness Analysis (COEA), *Proceedings of the IEEE 1994 National Aerospace and Electronics Conference,* volume 2, pp1188-1194.

[2] E.W.H. Keizer (1998), *Future Reduced Cost Combatant Study*, NATO report, 24 September 1998, NATO Restricted.

[3] M.H.M. Delmee (1998), *The Operational Value Model (OVM): a generic method to assess the operational value of a Future Reduced Cost Combatant (FRCC)*, TNO-report FEL-97-A352, NATO Confidential.

[4] T. L. Saaty (1990). How to make a decision: The Analytic Hierarchy Process, *European Journal of Operational Research, volume 48*, pp 9-26.

[5] V. Belton, T.J. Stewart, *Multiple criteria decision analysis. An integrated approach*, Boston, Mass.: Kluwer Academic 2002.

[6] H.W. Noordkamp, A.I. Martins Botto de Barros, B.J. Visser (2009), *A methodology for Simulation Based Performance Assessment*, TNO report TNO-DV 2009 A019 (in Dutch).

[7] *Methods and Models for Life Cycle Costing, Final report of Task Group SAS-054*, NATO RTO Technical report, June 2007.

[8] J. von Neumann, O. Morgenstern, T*heory of Games and Economic Behavior*, Princeton, NJ. Princeton University Press, 1944.

[9] P.C. *Utility Theory for Decision Making*, Huntington, NY. Robert E. Krieger Publishing Co., 1970.