SHIP AND NAVAL TECHNOLOGY TRADES-OFFS FOR SCIENCE AND TECHNOLOGY INVESTMENT PURPOSES

A Dissertation
Presented to
The Academic Faculty

by

Raffaele Gradini

In Partial Fulfillment
of the Requirements for the Degree
of Doctor of Philosophy in the
School of Aerospace Engineering

Georgia Institute of Technology
May 2022

SHIP AND NAVAL TECHNOLOGY TRADES-OFFS FOR SCIENCE AND TECHNOLOGY INVESTMENT PURPOSES

Approved by:

Dr. Dimitri N. Mavris, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Daniel P. Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Michael J. Steffens
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Mariel M. J. Borowitz
School of International Affairs
Georgia Institute of Technology

Gen. Philip M. Breedlove
School of International Affairs
Georgia Institute of Technology

Date Approved: January 7, 2022

The sea finds out everything you did wrong.

Francis Stokes

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Professor Mavris, without whose guidance I would not be the scholar I am. I was honored to be selected in 2018 to join the ASDL team, and throughout the years I have never regretted that decision. In the past four years I have faced many difficulties and obstacles, thank you for always taking my side. I hope in the future we will have further opportunities to collaborate.

To General Breedlove, who has instilled in me the passion for international cooperation and cooperative security. It is thanks to you that I have learnt a lot in the military domain, and on how NATO works. Our conversations were inspiring to me and are a source of guidance for my future endeavors.

To Dr. Steffens, I have truly appreciated working together for the past 4 years. I hope our paths will cross again in the future.

To my mom Chiara, a strong woman who in the years has been my model and source of inspiration. To my dad Nino, thank you for teaching me everything I know about the beauty of nature and how to appreciate it. To Gabriele, thank you for everything you do for me; you never had an easy job and yet you have always excelled.

To Beatrice, thank you for being there when the currents were strong and the tide was bringing me away, our common passion for the underwater world was one of the reasons why I picked the topic of this thesis 4 years ago. To this date I do not regret this decision.

To Cesare Augusto, thank you for being my best friend. Even if we were far, we never missed a conversation… We should have shared couple of beers more though.

To Austen and Polo, you have been my family in the US for 3 years. I cannot find the words to say how lucky I was to have you in my life, thank you for all the time, the laughs, the debates, and the good memories we shared.

To Flavia, Riccardo, Alex, Susanna, and Isa, thank you for being the lighthouse guiding me to port every time I come home. Rome would not be the same without you, and neither would I.

To Marc and Lea, thank you for being colleagues and friends throughout these years. We have spent countless hours working together, debating ideas, hypotheses, and assumptions. Needless to say, I would not have done it without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

A2/AD	Anti-Access Area Denial
AOR	Replenishment Tanker
ASCM	Anti-Ship Cruise Missile
C2	Command and Control
CBA	Capability Based Assessment
CDCM	Coastal Defense Cruise Missile
CESM	Communication Electronic Support Measure
CIWS	Close-In Weapon System
CONOPS	Concept of Operations
DAF	Defense Acquisition Framework
DD	Destroyer
DDG	Guided Missile Destroyer
DoE	Design of Experiment
FF	Frigate
FON	Freedom of Navigation
FONOP	Freedom of Navigation Operation
FPB	Fast Patrol attack Boat
IED	Improvised Explosive Devises
ISR	Intelligence Surveillance and Reconnaissance
JCB	Joint Capability Board
JCS	Joint Chiefs of Staff
LHD	Landing Helicopter Dock

MoE	Measure of Effectiveness
MOEA	Multi Objective Evolutionary Algorithms
MPA	Maritime Patrol Vessel
NATO	North Atlantic Treaty Organization
NAC	North Atlantic Council
NDS	National Defense Strategy
NMS	National Military Strategy
NSS	National Security Strategy
OPV	Off-shore Patrol Vessel
OSD	Office of the Secretary of Defense
PGM	Precise Guided Munition
RESM	Radar Electronic Support Measures
RoE	Rules of Engagement
SAM	Surface to Air Missile
SoS	Systems of Systems
SSK	Diesel Electric Submarine
SSM	Surface to Surface Missile
VUCA	Volatile, Uncertain, Complex, Ambiguous

SUMMARY

Long-term naval planning has always been a challenge, but in recent years the difficulty has increased. The degradation of the *security environment* is leading toward a more volatile, uncertain, complex, and ambiguous world, heavily affecting the quality of predictions needed in long-term defense technology investments. This work tackles the problem from the perspective of the maritime domain, with a new approach stemming from the state-of-the-art in the defense investment field. Moving away from classic methodologies that rely on well-defined assumptions, it is possible to find investment processes that are broad enough, yet concrete, to support decision making in naval technology trades for science and technology purposes. In fulfilling this objective, this work is divided in two main areas: identifying technological gaps in the security scenario and providing robust technology investment strategies to cover those gaps. The core of the first part is the capability of decomposing maritime assets using modern taxonomies, to map the impact of different technologies on ships. Once technologies are mapped, they can be traded inside assets, and assets inside fleets to quantitatively evaluate the overall fleet robustness. The first deliverable achieved through this process is called *Vulnerable Scenarios*, a list of possible conflict scenarios in which a tested fleet would consistently fail. The second deliverable is called *Robust Strategies* and is made of different technological investments to allow the studied fleet in succeeding the discovered Vulnerable Scenario. To find the first deliverable a large set of scenarios were simulated. The results of this simulation were analyzed using the Patient Rule Induction Method to isolate, among the large set of relevant cases, a subgroup of Vulnerable Scenarios. These

were identified by highlight commonalities on shared parameters and variables. Once the Vulnerable Scenarios were discovered, an ad-hoc adaptive response system using a "signpost and trigger" mechanism was used to identify different technologies on the ships studied that could enhance the overall robustness of the fleet. In identifying these technologies, the adaptive system was supported by different taxonomies in performing the different technological trades that allowed the algorithm to find Robust technology Strategies. The methodology was completed by a ranking system that was designed to firstly check all the Robust Strategies in all the scenarios of interest, and then to compare them against ranking metrics defined by decision makers.

To test the created methodology, several experiments were conducted across two use cases. The first use case, which involved an anti-submarine warfare (ASW) mission, was used to demonstrate the individual pieces employed in the creation of the methodology. The second use case, involving a large operation made of several tasks, was used to test the overall methodology as one. Both use cases were designed on the same original scenario created in collaboration with former generals and admirals of the US Air Force and the Italian Navy. The primary results of this experiments show that once Vulnerable Scenarios are discovered, it is possible to employ an iterative algorithm that recursively infuse new technologies into the fleet. This process is repeated until Robust Technology Strategies that can support the fleet are selected. The missions designed demonstrated the presence of gaps which had to be covered via technology investment showing how planners will have to account for new technologies to be able to succeed in future challenges.

The methodology created in this thesis provided an innovative way of enhancing the screening of maritime scenarios, reducing the leading time for investment decisions on naval technologies. In conclusion, the work done in this thesis helps in advancing the state of the art of methodologies used by planners when looking for Vulnerable Scenarios and for new technologies to invest on. Therefore, this thesis demonstrates that by employing the proposed methodology, Vulnerable Scenarios and relevant technologies can be identified in less time than by employing current methods. These efforts will support planners and decision makers in reacting faster to new emerging threats in unforeseen naval scenarios and, will enable them to identify in a rapid fashion in which areas more investments are needed.

CHAPTER 1. Introduction

History doesn't repeat itself, but it often rhymes.

Mark Twain

We live in years of great changes: countries once considered developing are now developed, the 2020 SARS-CoV-2 pandemic has shaken our societies with new, rapidly emerging problems and finally, the digital evolution is connecting us more and more while exposing our lives to new, and often subtle threats. The military world is not extraneous to any of this. On the contrary, many have pointed at a significant degradation of the *security environment*. According to Admiral Winnefeld, the *security environment* is one of the four pillars of an effective security strategy [1], together with *ends*, *means* and *ways*. Through the course of this chapter, the analysis of how the *security environment* is degrading will be presented in motivating the move toward a VUCA environment. To balance this degradation different *ends*, *means* and *ways* will have to change. To understand long-term planning this change must be studied not only considering *ends*, *means*, and *ways* in isolation, but also in the larger defense investment ecosystem.

1.1 The VUCA Environment

Armed forces of the western world are seeing a mutation of the *security environment*, which is becoming more Volatile, Uncertain, Complex, and Ambiguous, or as Bennis and Nanus coined in 1987, a VUCA environment. A deteriorated security environment has opened the doors to new challenges and threats to western countries' prosperity over the past 30 years. It opens new opportunities to state actors exploiting the over attention to terroristic groups, allowing them to quietly cover their technology gap. Finally, the opening of new and less regulated battlefield, as cyber, is an example of how the full international environment is transforming in a gray area where attacks are often way more subtle than what it used to be.

1.1.1 Volatility

In the volatile environment, problems arise at a much faster pace; they are brought by the state of dynamic instability present in a conflict when multiple interests, all pressing and all relevant, arise at the same time. Volatility requires people to be multitasking and it affects the decision-making process, as every choice must be taken with urgency and with potential drastic consequences [2].

The Middle East is an example of volatile environment which includes extreme polarization among the different parties. These countries can be divided in terms of their interests in the area and on their power. Countries like the US and Russia are in the region to expand their respective spheres of influence and to gain access to natural resources but others like Iran, Turkey, Syria, Israel and Saudi Arabia have been in conflicts for long periods, sometimes directly, sometimes by financing proxy groups to destabilize each other [3]. Historical, political, and religious matters have become so well integrated that they can create dichotomous beliefs within a region. Direct communications between Iran and Israel, or Iran and Saudi Arabia have been cut off in favor of propaganda statements and political messages. On the other hand, these states share deep roots. The melting pot of societies, the presence of powerful non-state actors, the weak governmental structure and territorial control of most states and the fast pace at which leadership changes has made of the Middle East the most volatile region in the world [4].

The intelligence environment can be also volatile. Information today are more fluid than in the past, and within a hyperconnected world it is easier to have leaks and stages of politicization and factionalism affecting the stability of a country. Looking again at the Middle East theater, leaks and politicization are among the main reasons why today the Iranian security sphere is considered volatile [5]. Often it is not possible to know if information were leaked on purpose or if it was by the action of external actors [6]. Volatility of information in this field is dangerous as it increases the uncertainty and the ambiguity of what is going on, simultaneously requiring an immediate response which could lead to an escalation of tensions.

1.1.2 Uncertainty

Uncertainty in the context of the VUCA environment means that experience alone is not enough anymore to advance. Nevertheless, decision making processes on the field still rely on what is known and what is not known. In defense investment processes, dealing with uncertainty means deciding whether an enemy has – or will have – a certain capability and what would be a contingency plan when facing the unexpected.

To prove how important it is to consider uncertainty in planning, an example would be to look at the role of the Ardennes forest during the German Blitzkrieg in 1940. Several French generals thought it was impossible for any heavy mechanized division to pass through the forest. As such, following this assumption they concentrated their defenses assuming the Ardennes was a natural, and impassable, hard border [7]. In this perspective the French plan was tailored on a fighting style similar to the one of WWI. This led to the construction of the Maginot Line: a series of fortification along the border with Germany meant to fight a trench war, not a war with highly mobile assets. The French commander-in-chief Maurice Gamelin was uncertain about the maneuverability and the mobility of the new Nazi mechanized divisions, which included both Panzer divisions and Light Armored divisions. On May 11^{th}, 1940 all the divisions managed to pass the forest and started pivoting toward the French fortifications which were then surrounded on 3 sides and cut off from the rest of the Maginot Line. For the French, this arguably meant the beginning of their surrender in WWII.

In the example it was shown how underestimating enemy's capabilities led to a catastrophic disaster which culminated with the loss of the Northern section of the defense line and the surrender of France few weeks later. This demonstrates how in an evolving world it is important to consider uncertainty.

1.1.3 Complexity

Complexity means that threats and opportunities, means and tools, cannot be considered individually but they must be considered in an interactive way. This idea was clearly expressed by former Defense Secretary, Jim Mattis, who in 2017 in a keynote speech at the Association of the United States Army's annual meeting, [8] described how different actors are contributing to the issue. The first element Mattis considered was a non-state one: terrorism in the Middle East. According to him, there are multiple states hiding behind their nation-state status while actively supporting terroristic groups. This practice has created destabilizing revolutionary regimes that increase their power by murdering and creating mayhem without any consideration for the human losses. The second element is ignoring international law, which happened in 2014 for the first time in Europe after WWII. In that occasion, a border was changed by force while a diplomatic option was still available. In annexing Crimea, Russia proved to be willing to discard Ukrainian's right to independently decide on economic, security and diplomatic matters. The third element is the presence of rogue states, actors like North Korea who are risking and threatening regional and global peace to reach global approval. The fourth element of the complex world which Secretary Mattis described is the role of space and cyberspace. A fifth element that Mattis does not quote, but that the author feels is relevant to mention is the presence of a near-peer power. After the end of WWII and until the end of the Cold War the world was divided between the American sphere of influence and the Soviet one. With the fall of the USSR the world transitioned toward a monopoly situation in which the US, thanks to a combination of military projection capabilities and diplomatic relationships, enlarged its sphere of influence. Today the situation is changing again, and many authors sustain that we have transitioned toward another duopoly situation in which, now, the two powers involved are the US and China [9]. All these five elements together are the reasons why today's threat environment is defined as complex. Military forces face a potential threat incoming from all possible axes, from Improvised Explosives Devices (IED) in Iraq, to maritime mines in the Strait of Hormuz in front of Iranian coasts, from cyberattacks to the nation critical infrastructure to new Chinese military bases on atolls in the Southern China sea.

In 2018 Mercier [10] described how NATO is gradually evolving to capture the evolutionary complexity of the security environment. He pointed out how NATO saw different phases: the Cold War time which was considered the age of collective defense (1949 – 1991), the post-Cold War time saw the Soviet Union and the Eastern bloc fall (1991 – 2001), the third phase started with the war on terrorism that began after 9/11, when NATO focused on projecting stability in the Middle East (2001 – 2014), the fourth and current phase began with Russo-Ukrainian crises of 2014 and has seen the resurgence of conventional threats supported by new hybrid threats (2014 – ongoing). It is interesting to note that all these phases have culminated in present day, in the Balkans. Mercer describes the region as the perfect melting pot of conditions that show how the threat environment is changing, and how the complexity is increasing. The Balkans were always in the sphere of influence of Russia, but now they also face an increase in radical Islamism, in conjunction with massive migrations and organized crime. Mercer also talks about how new technological means are more available to the public, empowering state and non-state actors with new threat domains and confrontational tools. The "threats domain" Mercer is referring to are cyber and informational, these are adding a new, complex layer as they allow the distribution of uncontrolled information to a much broader audience than was possible in the past. These information/disinformation campaigns are increasing their relevance as the veil between peacetime and crises becomes more and more blurred.

This multifaceted reality must be captured in new defense investment programs, avoiding focusing on one side of the problem only to find ourselves ready to fight a different war.

1.1.4 Ambiguity

All the elements seen so far contribute toward a general frustration and lack of clarity. Ambiguity leads to inefficiencies [11] and sometimes, if there is a lack of communication among the parties, to conflicts. An ambiguous world is a world where information can be interpreted in different manners.

Ambiguity in the context of the threat scenario can be seen when actions of a state or a non-state actor are not straightforward and contribute to clouding possible transitions between peacetime and crisis. The greatest risk of ambiguous threats is that they are not perceived as such, [12] driving the leadership toward either ignoring the problem or taking a wait-and-see approach. Ambiguity spreads well in what are commonly considered gray areas. These are where hybrid threats can proliferate in a holistic way transforming the civil society in a non-conventional battlefield. Gray area actions are shaped to avoid any military conflict, so while they are below any military response that could lead to a conventional war, they are still above what is considered fair competition among states and statecraft [13]. Examples of these actions were when invading Crimea and Donbass in 2014 Russia deployed unmarked troops asserting that those people were Crimean volunteers and local self-defense groups [14]. The Crimean “volunteers”, or Russian soldiers without patches, perfectly served the scope of annexing Crimea while mudding the water enough that a possible response from other countries had to face a fact accompli. Another similar use of ambiguous forces is the use of proxies. These have proven to be particularly effective in the Middle East, where states like Iran [3] have actively supported proxy groups to undermine the role of NATO in the region while avoiding any direct exposure. Iran particularly has been active in supporting several proxy groups such as Hezbollah in

Lebanon, Shia militias and paramilitary groups in Syria and Iraq, and the Houthis in Yemen.

On the other side of the globe, China leads ambiguous operations in the South China Sea. Here the communist state is transforming atolls into islands in international water while putting fortified military outposts on top of them [15]. By using this display of strength, China is trying to set up a solid claim on those waters which are currently disputed territory between China, the Philippines, Vietnam, Malaysia, Indonesia, and Brunei. China's strategy relies on military bases to support not only the official Coast Guard and Navy but also its enormous fishing fleets [15]. In at least two known scenarios the confrontation involved an unmatchable number of Chinese "civilian" boats surrounding islands versus few Filipino Coast Guard ships [16]. One after the other, ambiguous actions are pushing toward the final Chinese claim over the South China Sea as an Economic Exclusive Zone (EEZ), and while tensions are rising, so far, the strategy seems to be working [17].

1.1.5 Consequence of the VUCA environment

The consequences of this environment are directly reflected on the modernization plans, undertaken by different armed forces. This problem is particularly relevant in the Navy where assets designed 30 years ago must face new tasks which were not in the mind of naval architects. Asymmetric and symmetric challenges are arising rapidly, therefore while the US Armed Forces have been focused, in the past 20 years, on fighting terrorism developing *ways* and *means* for that scope, other countries have used the time to build up their conventional capabilities. This has led to a narrower technological gap between the US and other near-peer adversaries [1]. Answering this challenge is not the only objective of the American Armed Forces. In fact, in a world with two super powers, and several regional powers, the objective established by the DoD is to be able to fight a major war – and win it – while denying a secondary main adversary [22][23]. Having these two main objectives means envisioning a possible conflict on both the Atlantic and the Pacific scenario, or in the Arabian Sea and the Pacific Ocean.

Whichever scenario dictates the chessboard, it is a difficult problem to solve and as such for several years there has been a request to increase the number of naval fleets. For this reason, former Secretary of the Navy, Thomas Modly repeatedly said that a larger fleet is needed and desired, but that it will not happen without an increase in funding to avoid risking having the assets but not the men to run it, or as he said: a "hollow force" [20]. In 2020, with a fleet of 296 ships being still far from the 355 goal [21], a new goal seems to be approaching. The Pentagon is in fact now looking for a fleet of 530 ships by 2035 [22]. Differently from before though new *means* are expected to be on the table, and while it is still unknown how exactly the fleet is going to look like, several large unmanned platforms are being considered [23][24][25]. This possible shift in policies is interesting, as it would demonstrate an increased interest on new *means* to maintain the conventional technological gap wide enough.

1.2 Naval Domain

Among all the possible domains, this work focuses on the naval one. The reasons why this specific domain was chosen are two: its role in the security environment and the complexity of naval operations.

Regarding the security environment, Boyer [26] gives a wide description of the relevancy of the sea in our societies. The main point he touches is that we depend on the sea from food, to commerce, to border security, to natural resources. Therefore, he argues that a shift in the security environment will affect the seas and all the population living in coastal areas – which are more than 30% of US and European citizens. Boyer, therefore, pushes for a need in enhancing the naval response on both *means* and *ways* to ensure the resiliency of the military instrument of power at sea.

Complexity in the naval domain is shown by the types of assets, and their interactions, above and below water. As reported by the US Marine Corps [27], threats are increasing in complexity involving coast and oceanic assets above and underwater. This complexity is reflected in assets that are more adaptive, autonomous, scalable, efficient, fast, and lethal. Therefore, the naval domain right now is seeing a shift on what is in the water, with certain nations still employing Cold War assets, and others using unmanned vehicles. To be on top of this complexity, the USMC has stressed the need to identify possible future scenarios and to align the naval R&D to deliver the required capabilities at a much faster pace than previously [27].

Finally, from a system engineering point of view, fleets are interesting systems-of-systems. Each ship has unique capabilities and tasks, but when working together ships can exert more military pressure than the individual vessels. The complexity of the relationships among ships has been increasing in the past years as a consequence of the more variegate weapons and offensive measures available by adversaries. An example of this is the dichotomy effect that A2AD bubbles have on Vertical Launch Systems (VLS): on one hand ships have to allocate as many resources as possible to their protection, on the other if they survive with nothing in the VLS to attack, the operation is useless. As such, some have proposed creating VLS barges which will provide 256 or 512 launching tubes, but will fully depend on the fleet for protection, target illumination, support and so on. Modelling such fleet requires a much better understanding of what each ship can and cannot do, what the weaknesses and strongpoints are, and what are the benefits of bringing to the fight such powerful but fragile assets. More on future capabilities and technology trends is reported in Appendix A.

1.3 New ends, new ways, and new means

To respond to this degradation of the *security environment*, and to balance it, the other three legs (*ends*, *means* and *ways*) of the secure strategy must adapt. The question is how? How will the posture of the US change on the global stage? How would the mechanisms the US uses to communicate its foreign policies – militarily and diplomatically – be different? And finally, which assets and tools will the US use to signal its reaction.

1.3.1 New ends

A change of *ends*, the goal of the US posture on the global stage, is not likely to happen. After the fall of the Soviet Union, the US had the occasion to expand its interests all over the world almost undisputed [28]. Even with China currently trying to rise as a superpower it is unlikely that the US will be ready to give up its posture of dominance to adapt to a new equilibrium. In 2014 the Deputy Secretary of Defense, Robert Work, clearly stated that the posture the US has adopted is forward oriented to intersect, project and protect national interests. In the same interview Deputy Secretary Work stated that the US recognizes only itself as a global power and that it will keep using its *means* and *ways* to keep this equilibrium [29]. Several analysts and, later, DoD Secretary Esper, have remarked how the US is indeed keeping its posture against China by doing several Freedom of Navigation Operations (FONOPS) in the area. In May 2020 to reiterate the position of the US on the matter Secretary Esper stated the following:

"I don't know what the Chinese meant by that hollow statement about American carriers being there by the pleasure of the PLA [People's Liberation Army]. *Look, American aircraft carriers have been in the South China Sea and the Indo Pacific since World War II, and we will continue to be there, and we're not going to be stopped by anybody. We're going to sail, fly and operate where international law allows"* [30]

1.3.2 New ways

Ways are the strategic and operational concepts employed by nations to secure and maintain national interests.

The US military strategy is founded on the assumption that to defend American interests the US must be able to offensively project its military power to strike against adversaries. This attitude has been supported in the years by several conflicts which have seen a clear advantage on the attacker side, the Falkland war being a clear example of this. But this might be mutating with the new environment, and the attacker may no longer have the advantage it used to.

A signal of this shift was seen in the 2018 National Defense Strategy [31] by Secretary Mattis who reported the need and urgency to act to preserve the current strategic advantage of the US. Mattis stated that without any response to China and Russia's defensive advancements, they would not be able to cover the technological gap and gain the ability to effectively deter an act of aggression or any coercive strategy around their regions. Years of the US focusing on the war against terror has given the advantage to other countries – China and Russia – which were able to develop new conventional *means* to block American current strategies (*ways*); de facto starting to mine American overall secure *ends*. Once the gap will be fully covered, analysts call for an increased risk of escalations as many of the operations carried today to enforce the US foreign strategies will no longer be possible. Among the different pivotal operations that will have to change there are FONOPs.

The UN Convention on the Law of the Sea (UNCLOS) is an international agreement under which nations divided and shared the rights and responsibilities of the waters around the globe. Among the different regulations, one of the most important principles is the Freedom of Navigation (FON). FON is the norm which states that innocent passage through territorial water must be granted to all ships, civilian and military. To ensure this principle, many countries have organized different types of operations, in the US these are called FONOPs. These operations involve a close passage, within the twelve mile territorial water limit, of a military vessel to either ensure *innocent passage* is granted or to conduct military operations showing that the claim on territorial waters does not fall within the UNCLOS [32].

But freedom of navigation is currently under threat. With the fielding of supersonic YJ-12B anti-ship cruise missiles and HQ-9B anti-air missiles on the military installations of Subi, Fiery Cross and Mischief reefs [33] China is strengthening its posture in the region. These claims were reiterated in 2018, when an US P-8A Poseidon patrolling the area over Mischief Reef was warned by radio that it had violated China's sovereignty infringing its security and its territorial rights [34]. While at the time, the plane kept its course and finished the mission, the subsequent deployment of HQ-9B anti-air missiles on that reef might suggest a shift in the willingness of Chinese authorities to have American planes flying over their military bases.

The US must find new *ways* that need to be coherent with two pillars of deterrence strategy: *imposing costs* and *denying objectives*. In the context of the current security scenario, the goal remains to deter adversaries from gaining too much advantage which could lead to a conflict escalation due to a shift in the political equilibrium.

1.3.2.1 Imposing Costs

Imposing costs is a strategy that, if done correctly, imposes a high price to achieve something that would otherwise be much easier to an adversary. If, for instance, China would become more aggressive in the use of their bases in the Southern China sea, then it will be imposing costs on the US for performing a FONOP mission. In other words, the US would need a fleet which could safely and successfully pass within twelve miles of an aggressive military base instead of the single DDG, normally used in these kind of missions [35]. The economic cost of using a full fleet instead of a single ship to perform a FONOP mission would be the *imposed cost*.

Imposed costs though do not have to be only military. They could be used as part of other different *ways* to implement a foreign policy. An economic example could be the disruption of a supply line to damage the Chinese's Belt and Road Initiative, for instance by stopping commercial ships or in more extreme cases by mining some of the key infrastructure. This would force China either to spend money repairing the infrastructure or to find alternative, and less optimal, solutions reducing the return of investment [1].

1.3.2.2 Denying Objectives

Conversely to *imposing costs*, the role of defense is moving toward anti-access resources to target exposed forces in such a way that even if the attacker tries to achieve an objective via a hostile action, this is countered, and the objective denied.

Denying objectives is critical in two possible scenarios. The first and most obvious one is the defense of the homeland, by having strong coastal defense systems it is possible to create an Anti-Access Area-Denial (A2/AD) bubble across all the different domains. These defense systems work by superimposing different capabilities like Surface-to-Air Missiles (SAM), Surface-to-Surface Missiles (SSM) and Coastal Defense Cruise Missile (CDCM) [36]. The second case is the denial of objectives to alien countries in foreign territories. In this case the full suite of SAM, SSM and CDCM might not be available, or it may be limited to what is available in a close-by base. The presence of an underwater force capable of piercing through A2/AD defenses helps in denying objectives even when most platforms are not available.

1.3.2.3 An example of a new way: Anti-Access/Area-Denial Capabilities

Objective denial and *imposing costs* can be seen together in the A2/AD capabilities deployed by Russia in the Kaliningrad Oblast [37]. In the past years Russia has built up, in the region, a complex array of weapon systems to interdict any access to the area and its surroundings by any aggressor. The Russian A2/AD system includes several SAMs as the SA-21, Su-24 aircraft ready to be scrambled, the Monolit-B self-propelled coastal radar targeting system which has over-the-horizon precision detecting, classification, and determination. Several Anti-Ship Cruising Missiles (ASCM), and mobile CDCM platforms are also in place, complementing the air defense systems with stand-off anti-ship missiles and land attack missiles. The geography of the Baltic Sea allows for very few waterways to enter in the area, therefore even a small number of Kilo-class submarines are an efficient patrol. Moreover, these waterways can be transformed in to checkpoints and minefields to block any unwanted navy from entering in the area if needed.

What Schmidt suggests in his overview on countering A2/AD [38] is the need to counter two specific objectives to at least partially disrupt the A2/AD system. The first one is the need to counter Intelligence, Surveillance and Reconnaissance (ISR) by jamming the sensors (radars, sonars, cameras, etc.) of ISR assets. Jammers can not only deny ISR capabilities needed for precision targeting in most strikers, but they can also affect Command and Control capabilities disrupting communications. The second objective is to destroy Precision Guided Munition (PGM) platforms, this will have to be done by attacking from outside the A2/AD area. According to Schmidt there are two *ways* to counter A2/AD: Inside-Out or Outside-In. Inside-Out implies having a large technology advantage that allows for a quick and high-intensity conflict where the goal is to suppress the A2/AD to allow the arrival of support forces. The second strategy, Outside-In, is focused on dismantling, either by absorbing or by destroying, A2/AD capabilities layer by layer. This second approach can be lengthy, and it can lead to an escalation due to the increased attrition rate and is often not sustainable due to mission fatigue.

This leads to the third point: *means*. What are the new capabilities that nations should invest on to be able to still achieve the desired *ends?*

1.3.3 New means

Means are the tools in which the nation invests, they are assets, technologies and capabilities needed to perform missions and to protect the grand strategy of the country. Increasing resources in means is needed to compensate the growing imbalance among the secure strategy’s four variables. This thesis focuses on means, particularly maritime *means*. Although, in the modern world no asset can be considered by itself, therefore while still limiting the discussion to key concepts also the other four domains (Space, Cyberspace [39], Air and Ground) will be discussed.

When talking about investing in new means, the problem can be seen in at least three possible ways: investing in completely new assets, in more assets, in better assets, or in ways of using the assets.

Acquiring new assets is the first and more obvious consequence of the need for new *means.* New assets require an extremely long time to be designed, assembled, tested, and fielded. At the same time due to the complexity of the system only few new technologies are usually infused to reduce the risk of re-design or overextended testing phases. This has been seen in the development of the DDG-1000 [18][40] which after almost 15 years of design had to go through a massive re-design phase. Moreover, given the delays due to the multitude of new technologies onboard, the price of each unit increased by almost 50 percent, to a total of $13.8B for the first three ships [40].

Acquiring more assets is another viable option. While many navies claim that many assets are multi mission, each asset can still only perform one task at the time, therefore in scenarios where complex operations are happening, like in the A2/AD described before, more assets can lead to a strategic advantage. It must be mentioned though that more assets will not always bring more robustness to the fleet, on the contrary in some cases buying older assets will compromise operations. For instance, a newer and quieter ship will have better results in trying to locate a submarine than two older and noisier ones together.

The final option when having new means is to improve upon assets already owned. Infusing technologies in assets is not always possible as some of the technologies must be embedded from the early design, like stealth capabilities. Ships which are considered stealth have in fact specific features that are usually introduced in new design, and it is not possible to modify the geometry of the hull to make it stealthier after it is produced and delivered. On the other hand, it is possible to retrofit a new radar or sonar which will increase the detection capabilities of the ship. These are considered incremental technologies that can be added or changed throughout the operational life of the ship. Upgradability is relevant to the point that modern designs usually include more space allocated to for future instruments to be mounted on board.

1.4 The investment Process

To invest in new *means* the US has an extensive set of options depending on the entity of the contract, urgency, and the type of asset or service needed. The main process that supports defense acquisition is called Joint Capabilities Integration and Development System (JCIDS) [41]. This process is designed to validate requirements on warfighting assets and support the identification of such requirements and capabilities through contribution to the NDS and NMS. Throughout the process, operational analysis and operation costs are compared to balance the "return of investment" that acquired assets will bring as enabled capabilities and warfighting options. A partial representation of JCIDS is given in Figure 1 where greater emphasis has been given to the first part of the diagram: Capability Based Assessment (CBA).

CBA is focused on assessing and recommending whether material solutions are sufficient and if non-material solutions will satisfy the need. This analysis is focused on understanding and evaluating the gap to be covered through two main approaches: top-down and bottom-up. The top- down approach comes from the request of developing specific capabilities for high level stakeholders like Joint Chiefs of Staff (JCS), the Office of the Secretary of Defense (OSD) or the Joint Capability Board (JCB) [42]. The bottom-up approach can be initiated by Regional Commanders which elicit specific capability development via a Formalized Asset Request [43]. This leads to a study of the different solutions available, or to be developed, to cover the problem elicited by the commanders.

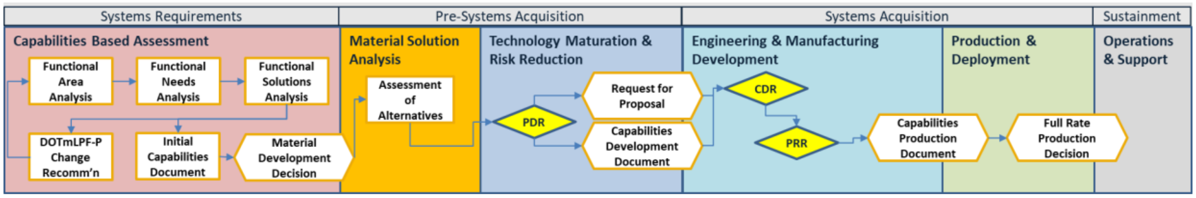


Figure 1: Major Capability Acquisition process adapted from [44]

Independently of how the process is initiated, it should then be a fast study. The goals of the study are to identify the required capabilities, identify the options available, evaluate risks and gaps, prioritize gaps, and identify viable potential solutions. The problem though, a planning gap, appears when the top-down and bottom-up approaches must be matched. This issue is exacerbated by the heavy reliance on subject matter experts (SMEs) on both top-down and bottom-up approaches. The planning gap today is not really covered, attempts to do so are present in similar processes to the JCID used by specific branches of the Armed Forces. This is the case of USAF Technology Development and Investment Process Figure 2, where an iterative method was developed to cover the mismatch between bottom-up and top-down approaches. Nevertheless, the reliance of CBA on subject matter experts remained. This led to a limited view on the evolution of the security scenarios, as SMEs relies on their experience to forecast expected scenarios and technology needs. The reliance on SMEs, is therefore a limiting factor that affect how planning is conducted, and that has an impact on the preparedness, from the technological point of view, of Armed Forces in future conflicts. Being able to move away from using SMEs for scenario forecasting will be one of the objectives of this thesis. In the conclusion chapter, it will be demonstrated how obtaining the same deliverables of a CBA through a process that does not use SMEs for discovering scenarios, assessing risks, and identifying technological gaps is possible.

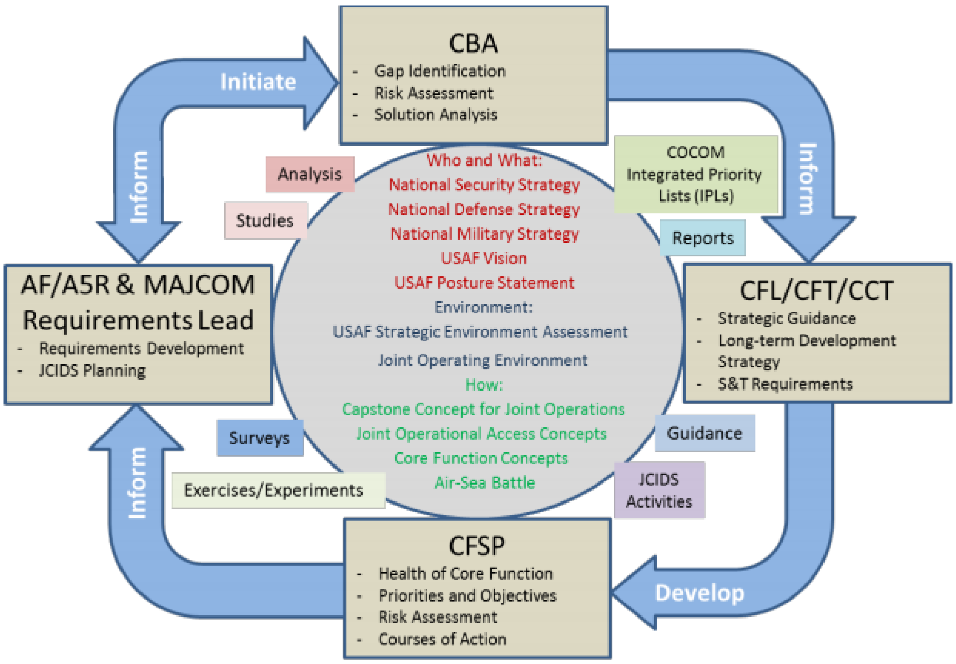


Figure 2: USAF Technology Development and Investment Process.

1.5 Planning gap

Many have criticized defense planning in the United States, saying that there is too wide of a gap between bottom-up approaches, such as resource-based planning, and top-down threat- and capability-based planning [45]. This gap appears more evident when the information used by the relative stakeholders are matched. The top-down approaches initiated by JCS, JCB or OSD are based on a much broader picture, which includes assumptions from the global scenario [46]. Focusing on operational details and planning of several fleets deployed around the globe. Moreover, because the problem in this approach is seen from the top, there is a risk of achieving ill-defined solutions that do not cover all problems in the field. On the contrary, bottom-up approaches focus on specific problems present in one theater that might not be present in others. In this approach the different commanders put together a set of requirements based on their current operations and forecast the threats they might face in the future. The risk in taking this approach is hyper-focusing to focus too much on answering one specific problem without keeping in mind that those threats and assumptions may not be valid in a different timeframe, or in a different location.

In general, experts suggest that these methods are no longer adequate for mid- and long-term science and technology investments [47]. The gap shows how current investment strategies could lead to either single point solutions, in which uncertainty is not considered, or solutions that are too focused on abstract threats. In the end, the risk is formulating a plan which leads to impractical, not credible, and short-sighted solutions.

1.5.1 Covering the Planning gap

This gap must be covered in a quantifiable way. It is important to start analyzing the mission from the beginning, dividing it into several tasks performed by the assets. Without losing generality, it is possible to say that those tasks should be characterized by quantitative Measure of Effectiveness (MoE). MoE can therefore be used not only to describe tasks, but also to evaluate how, and if, different *means* will satisfy these tasks in a quantitative and trackable way. Having quantifiable MoE is key to defining the requirements each task will have to satisfy throughout the entire mission. By using MoE it is also possible to track the evolution of requirements' satisfaction during the mission, analyzing critical phases that could lead to mission failure.

Quantitative MoE are the key for an investment strategy that is not focused on any specific technology but that is driven by requirements. From this statement emerges the implicit assumption that we should be able to trade quantitative metrics describing the technologies to evaluate the full investment strategy.

1.6 Research Objective

When looking at what has been presented so far, it is possible to see that many support the conclusions that the degradation of the *security environment* is a broad issue and that it should be answered by an increase in defense investments. This degradation is fostered by state and non-state actors which have been pushing for a more volatile, uncertain, complex, and ambiguous environment. Because of this shift toward a VUCA environment, understanding the needs of what should be invested in, is clouded [48] and, as Winnefeld as said, can lead toward acquiring capabilities for the wrong type of conflict. To rebalance the table, the United States must decide if it wants to change its current strategic goals – *ends* – or if it wants to maintain the current military posture. This decision is the main driver for what will be the different strategies – *ways* – and assets – *means* – it will invest in the future. While recognizing the dual aspect of the problem, this thesis focuses on the capabilities of the *means*, while keeping the *ways* constant on the friendly side. On the enemy side, the degradation of the *security environment* has pushed toward the adoption of new *means* and *ways*. As such, new capabilities will have to be countered on future battlefields.

The decision of investing in new capabilities, being new assets or new technology, cannot be separated from the legal and institutional aspect in which that decision is taken. As such, part of the Capability Based Assessment (CBA) has been reported. In the CBA, different gaps are analyzed and solutions are proposed by two groups of stakeholders: those acting in a top-down fashion like JCS, OSD and JCD, and those proposing investment through a bottom-up stream like Regional Commander. The gap between the two views must be addressed to align long-term investments on capabilities as seen from Washington to what is needed, and will be needed, by the men on the field. In covering this gap, the focus is to link in a quantitative way the technologies of interest for investment with mission requirements. Missions are therefore decomposed in different general tasks which are evaluated by different MoE. This decomposition, together with the use of quantitative MoE, allows for the evaluation of the investment strategy being driven by a requirement pull rather than a technology push. To address the aforementioned problems of the security environment and those of the investment strategy we define the following research objective:

Research Objective: Develop a methodology to support concurrent trades-offs among naval assets and technologies, to assist investments on new long-term maritime technologies.

1.7 Proposed Approach

To address the research objective a deep literature review was conducted. Before going in the details of each area, it is important to stop one moment and to appreciate the logic that was used in building this doctoral thesis. The logical approach showed in Figure 3 will be used as a map to guide the reader in how this work was shaped, so that he or she will know what to expect next and what are the logical links. It is worth mentioning that in writing this work it was decided to split the thesis in two areas: one theoretical and one experimental. Therefore, the reader will find all the theory bits up to Chapter 5 and all the experiments from Chapter 6 onward. This way, the reader can appreciate the theoretical build up all in one go, and can find all the experiments in the second part without jumping from theory to practice multiple times.

1.7.1 Logical Approach

In Figure 3, it is possible to see the logic used to address the Top-Level Question on how ships and naval technologies trades-off are conducted for science and technology investment purposes. To answer this question, it is important to understand on one hand how new means and ways are chosen when dealing with degraded operational environments, and on the other, what are the gaps in current planning practices. To investigate these two areas the work has been divided in 3 streams: taxonomies in modern naval systems-of-systems, modelling and scenario discovery techniques for decision making, and technology selection practices in investment methodologies.

The work on taxonomies was driven by an interest for an increased understanding in how components of a complex system interact, and how can the links be modelled. The alternatives studied looked at different type of taxonomies, ranging from vertical ones to hybrid evolutionary ones. Part of the work was also devolved to validating the quality of inputs, to ensure outputs' quality. The workstream on taxonomies ended with the identification that as long as the taxonomy chosen is hybrid enough that will coherently describe the relationship inter-SoS, then it will be possible to use it to comprehensively describe interactions among SoS.

The second stream of work focused on modelling and on scenario discovery techniques for decision making. This workstream was focused on answering the need to aggregate existing approaches to quantitively address future threat scenarios for long term planning. The first segment of this stream was dedicated to selecting the different modelling techniques that could better simulate naval future scenarios. After the modelling decision fell on agent-based, it was decided to investigate which technique could be used to find Vulnerable Scenarios. Due to the type of problem, and to benefit from the work done by RAND in similar areas, it was decided to use Scenario Discovery and in particular the Patient Rule Induction Method (PRIM). Unfortunately though, this was never applied to quantitative naval scenario discovery and therefore more work was needed to adapt it. Efforts were made to augment PRIM with Principal Component Analysis, to account for those situations where dataset had highly irregular geometries. The work on this branch of the thesis culminated with the drafting of the hypothesis that PRIM, ABM and PCA can identify Vulnerable Scenarios in a quicker manner than state-of-the-art methodologies. The hypothesis was then tested in a later experiment in Chapter 6.

The third workstream brood over different ways technologies can be selected to enhance fleets' long-term robustness. Among the different options, given the size and the type of problem, it was decided to use a hybrid approach which included both a DoE and iterative algorithm. Unfortunately though, when looking at selecting hybrid approaches that employed both DoE and iterative algorithms, it appeared that none of what was available satisfied the requirements on the method being quantitative, adaptive and able to handle many variables. Therefore, a new method was needed to enhance state-of-the-art methodologies with modern analysis tools, meant to address quantitative needs. This newly created method was hypothesized to be able to adaptively find sets of technologies that would increase fleets' success rate in naval scenarios. In testing this hypothesis, it emerged the need to verify that all different technologies tested are positive monotonic. This ensures that the infused technologies are not detrimental to the fleet but that they either bring only positive benefits, or that the negative effects are compensated. Finally, this branch looks at ways to reduce the number of fleets tested in an effort to reduce duplications and computational powers.

All the 3 workstreams are united in the last part of the diagram where the methodology is created, and the overarching hypothesis stated. At this stage the equivalency conjecture is drafted, this is needed to compare the results of the proposed methodology with what is created in a CBA. Finally, the proposed methodology is tested in a comprehensive experiment, using a different use case from what was used to demonstrate the hypotheses in the three workstreams. The methodology created will be summarized in the next section and described in detail in Chapter 5.

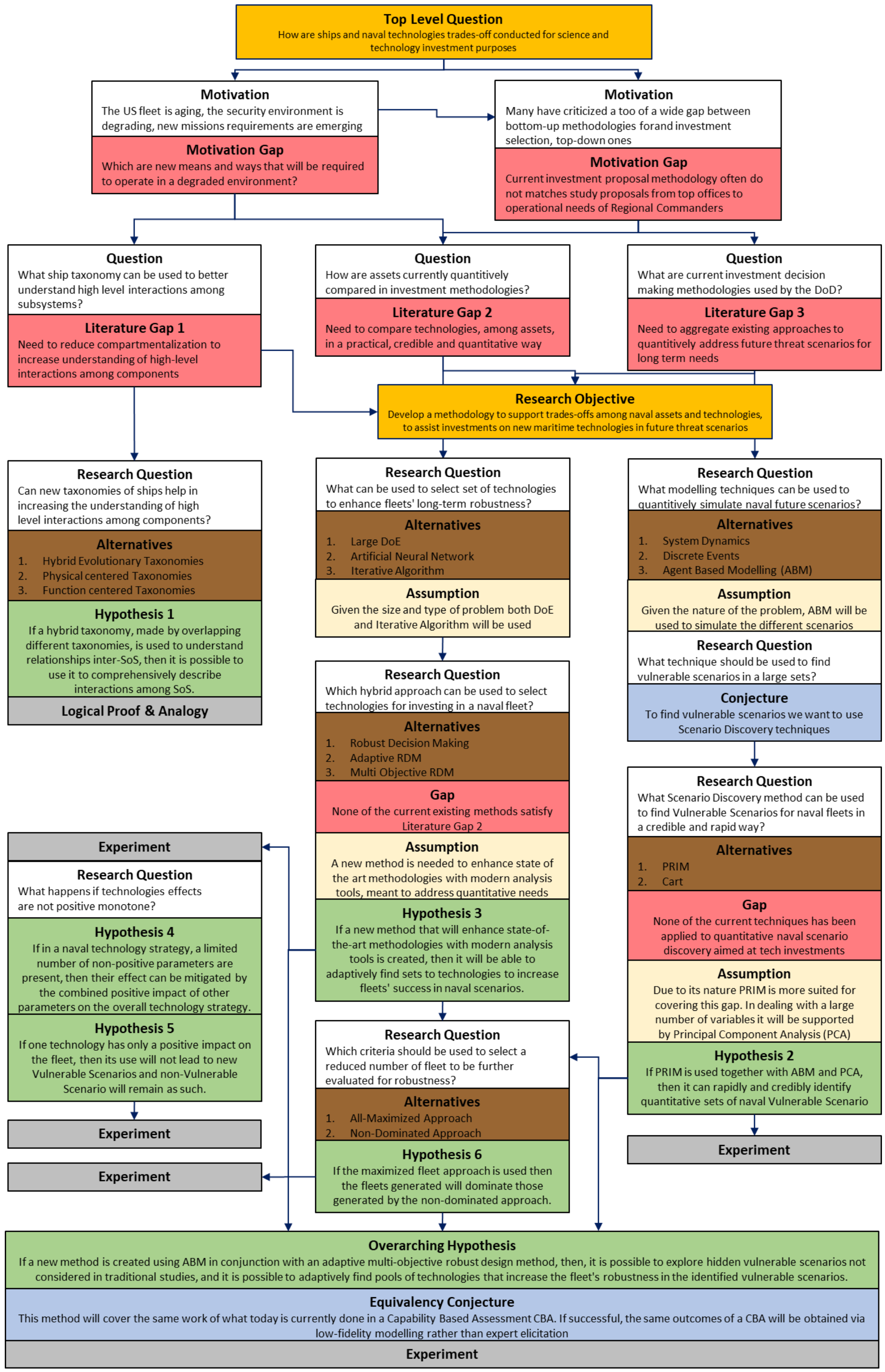


Figure 3: Logical diagram of the thesis development

1.7.2 Proposed Methodology

Following the logical approach descripted in the previous section, it is important to highlight the main parts of the methodology created.

The methodology was divided in three parts. The first one is dedicated to creating the framework for the whole problem and to formulating the problem from the scenario point of view and to identify Vulnerable Scenarios through the Scenario Discovery step. Taxonomies are also used in this step to identify a series of parameters and technologies of interest that will guide Part 2.

In Part 2 the first step is centred on verifying the positive monotony of the variables of interest to ensure that no new Vulnerable Scenarios are created. Once the different technologies are cleared, they are used in the iterative step as means to improve the studied fleet's ability to overcome the Vulnerable Scenario it was tested in. When a fleet is successful, it is outputted as Evolved Fleet

In Part 3 Evolved Fleets are reduced in numbers so that from each Vulnerable Scenario only one fleet is created. Following this operation, all the evolved fleets are tested in all the scenarios to identify those that are successful across the widest set of scenarios. At the end of the simulations, different ranking criteria can be used to rank the fleets depending on their success level, number of modifications and cost of investments. The ranked fleets found are called Robust Fleets as they proved to be successful in a wide variety of scenarios.

The goal of this methodology is to create the same deliverables of a Capability Based Assessment (CBA), with the idea that a possible end user for this work would a planning section in the DoD looking at the prioritization of technology investments.

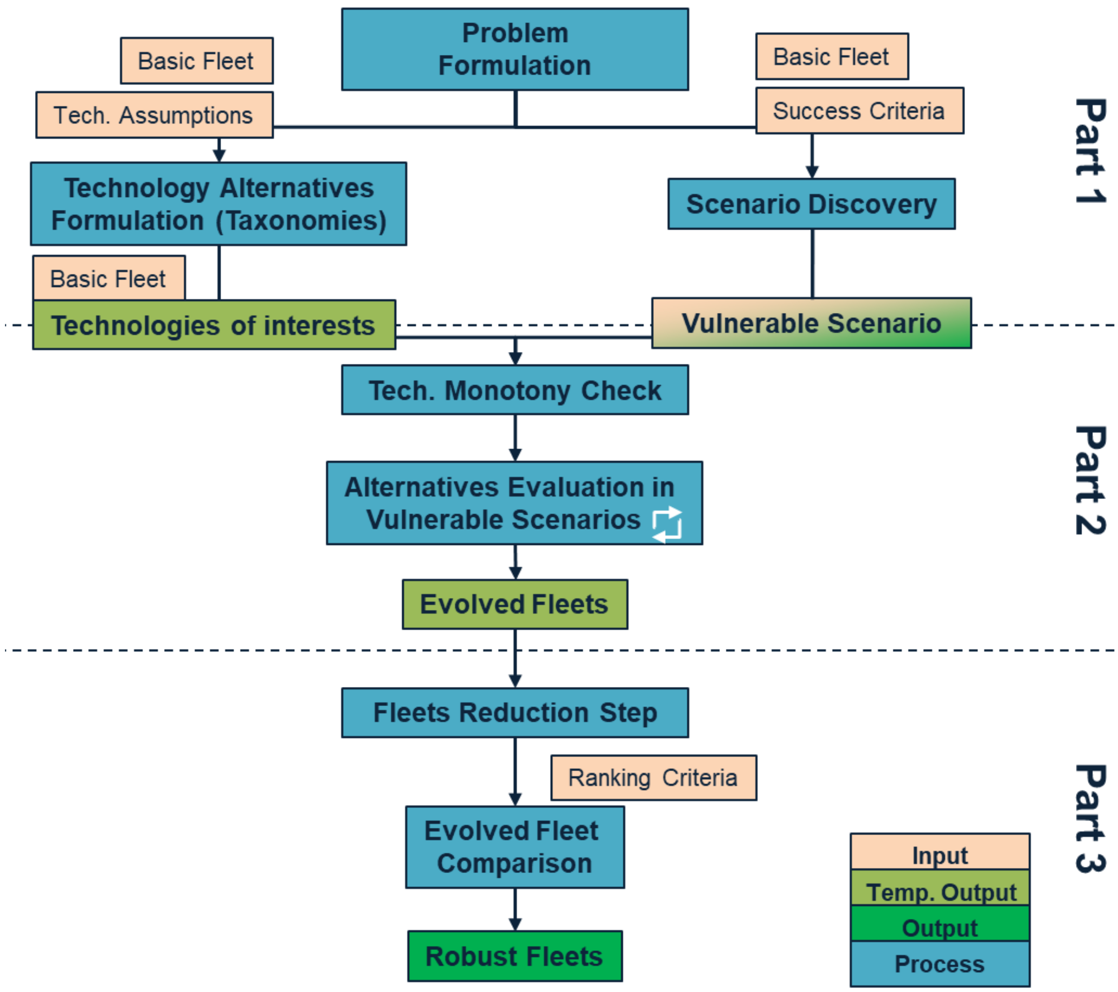


Figure 4: Proposed methodology used to achieve the research objective.

1.7.3 Proposed Scenario

Throughout the course of this thesis multiple experiments are conducted. While two use cases were chosen, one for the buildup of the methodology and one for the demonstration of the complete methodology, they both stem from the same scenario. This scenario was created and validated in the context of the NATO STO AVT RTG 317, where admirals and expert from naval and aerospace industries provided inputs and suggestions on the type of assets and on the development of the scenario itself.

The scenario created is set in a fictitious archipelago named Relas, here there are 4 countries: Relis – considered a NATO ally, Asyr – considered a NATO adversary, and Prucy and Turim both neutral but respectively aligned with Relis and Asyr. Relis and Asyr share the biggest island of the archipelago while Turim is located on the north and Prucy on the west. Figure 5 shows the archipelago composition and the respective territorial waters. From this set up two different use cases were created, the first one focused on an ASW mission and used in demonstrating the hypotheses individually, the second focused on a more complex and comprehensive operation, used to demonstrate the whole methodology.

In the first use case a fleet of two submarines is patrolling the area south of Prucy. One frigate is sent to try to locate the submarines and, if the submarines attack, to neutralize them. In this use case the frigate can either use its onboard sonar or the sonobuoys deployed by its helicopters. The mission is successful if the submarines are found in a certain time frame and if both the frigate and the helicopter are still alive. This task was chosen to test the various hypotheses because, on one hand it is simple enough that behaviors are predictable, and on the other it is complex enough to allow finding interesting cases where different technologies are needed. This use case is used in all the experiments from Chapter 6 to Chapter 9.

The second use case looks at a complex mission in which a fleet from NATO is sent to block Asyr from invading Relis. The mission is divided in 3 tasks (AAW, SED and NEO) which will have to be executed in a limited timeframe and without losing any asset. While the NATO fleet is approaching, Asyr employs different tactics and means to try to stop it. The goal of the NATO fleet is to reach the gulf in front of Relis' only port to evacuate civilians. Before doing that, the fleet has to successfully survive attacks from Asyr Air Force, and from a series of coastal defense stations hidden on Asyr's and on Prucy's territories. This use case is used to demonstrate the overall methodology in Chapter 10.

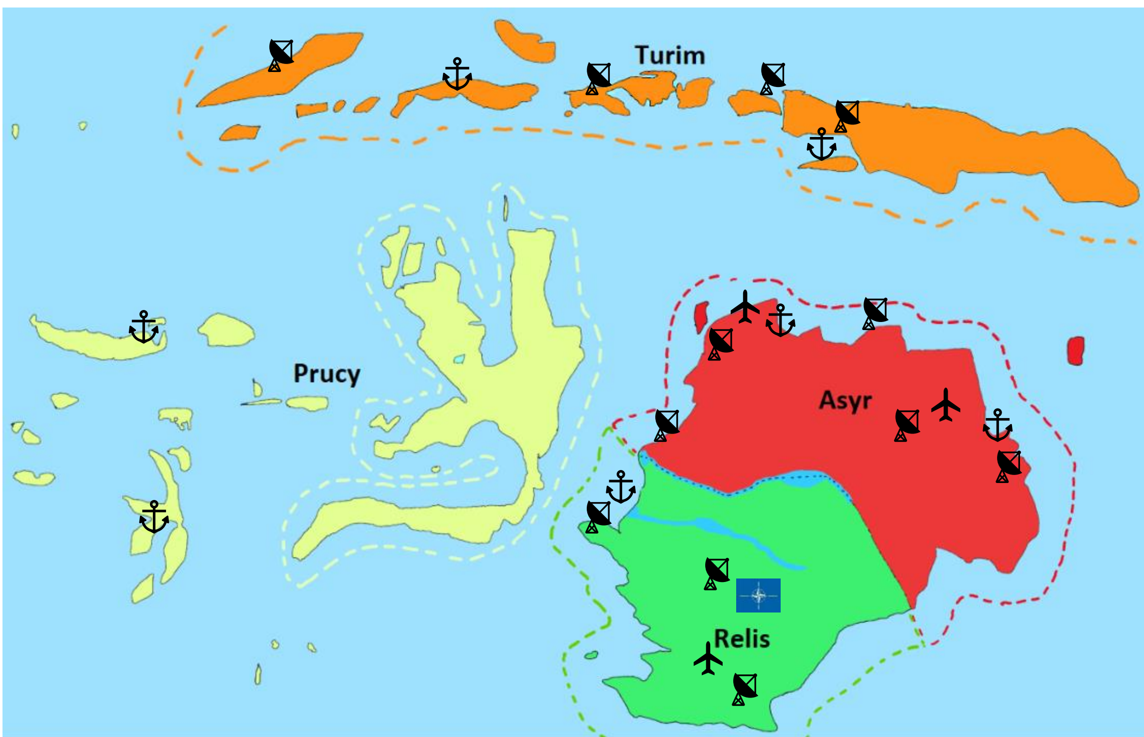


Figure 5: Map of the scenario location

CHAPTER 2. Taxonomy in modern naval Systems-of-Systems

2.1 Why taxonomies are relevant

Taxonomy is defined as the branch of science concerned with classification. This has played an important role in the military world since the late 18th century, when the military theorist Carl von Clausewitz pushed for a better understanding of systems and situations "in the blink of an eye". Von Clausewitz was interested in immediately grasping the consequence of performing a certain maneuver, or of having a certain piece of equipment to anticipate its role in the battlefield. The idea behind this was to organize different concepts in an ontology within a specific domain [49]. Von Clausewitz was interested in taxonomies not only from the tactical point of view but also from the engineering one, as such he was a pioneer in the field.

The reason why taxonomies still play an important role today is that they enable different ways of looking at the same problem. Focusing on one area or another, taxonomies expose unique properties of the system that would not be visible otherwise. Taxonomies become particularly relevant when introducing new technologies in already existing systems, as they allow a high-level understanding of the interactions between the new and the old system. Today, there are different taxonomies used in the field to improve the understanding of the system itself. The most common ones are traditional hierarchical taxonomies, as functional ones. All these play important roles in describing systems, and in decomposing them to basic elements which can be compared.

In this context, taxonomies will be fundamental to describe how ships can be decomposed and to understand the consequences of mounting specific technologies on the full ship. The chosen taxonomies will support the decomposition of assets of interest in the technology selection segment of this thesis, by doing so the impact of the chosen technologies on the fleet will be highlighted and comparisons made possible. Before investigating new taxonomies though, it is important to be able to answer the following research question focused

Motivation Question 1: What ship taxonomy can be used to better understand high level interactions among subsystems?

2.2 Currently used taxonomies

There are many different taxonomies currently used in the military world. From more traditional hierarchical ones, to more advance horizontally structured ones, their role is always the same: depict the system and show connections among its subsystems.

2.2.1 Traditional Taxonomies

In traditional taxonomies different subsystems are arranged hierarchically across different levels. Common traditional taxonomies are *Functional Taxonomies*. Usually, these types of taxonomies do not look at the broad system or system-of-system, but instead they focus on a particular area of the problem (e.g. the taxonomy for a C4ISR network) [50].

These types of hierarchical taxonomies are useful to decompose systems and systems-of-systems with a central authority, but due to their fixed vertical arrangement they fail at reproducing chaotic and horizontal systems [51]. Due to the vertical structure and the rigidity in connecting functions and components, systems appeared rigidly framed. As such, many have said that there is a lack of communication among SoS' components, limiting decision makers understanding of the SoS. The direct consequence is also a reduction in model precision, as modelled interactions struggle to reproduce what happened in the real world.

Observation 1: Traditional taxonomies are hierarchical and compartmentalized blocking the understanding of interactions between components present in different subsystems.

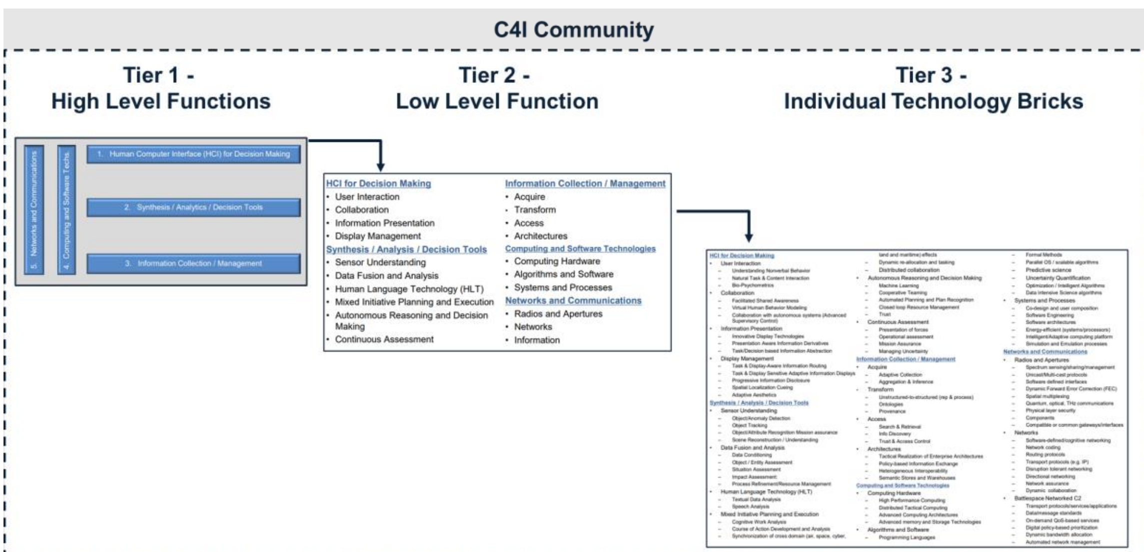


Figure 6 Hierarchical taxonomy of the C4I community [52]

2.2.2 *Task-Centric Taxonomies*

Task-Centric approaches are useful to organize technologies in relation of to the tasks they should perform. To achieve this, tasks are divided in families and subfamilies following the different needs of the system. By doing so, the operational scenario is directly linked to the technologies on board of the modelled asset. Modelling and simulations exploit this link to increase results' quality.

While this taxonomy is still hierarchical, it mitigates the excessive compartmentalization present in traditional taxonomies by using more fluid connections among systems' components. Unfortunately, by creating a fluid network of information, the readability and the understandability of the system are diminished. This is expected as these taxonomies focus on highlighting the different components participating in a specific task, independently from the physical or functional subsystem in which they are located.

Observation 2: Task-Centric taxonomies show how functions connect to different physical components in a complex way. This reinforces the understanding of which component is involved in a single task, while unfortunately it reduces the overall understanding of the system.

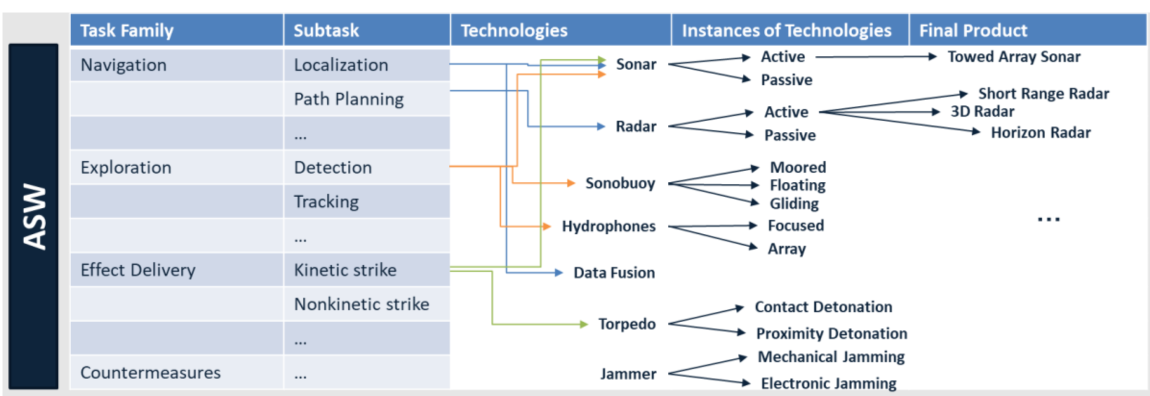


Figure 7: Task-Centric taxonomy adapted from [53]

2.2.3 Capability-Centric and Classification Taxonomies

Capability-Centric taxonomies are useful in decision making as they are used to decompose desired capabilities up to reaching the physical components that enable those capabilities. Among the different taxonomies currently used for this task, the most important one is the OODA loop [53].

Classification taxonomies are used in the military field to organize systems-of-systems and they tend to look at the problem from three different angles: Acquisition, Operation and Domain [50].

Acquisitions are classified as dedicated or virtual. The first refers to the full SoS when acquired in one block. The latter to assembled SoS with components not originally designed for that purpose, C2 systems are usually a good example of this.

An operational classification is observed when looking at relationships among assets during missions. A *chaotic SoS* is defined when assets, during missions, are fully independent and there is no central authority. In contrast, in a *directed SoS* there is a central authority commanding over the different assets deployed. In *collaborative SoS*, decisions are collaboratively taken by the different assets with little, or no, influence from a central authority.

Domains can be also used for classifying different SoS being whether physical, conceptual (i.e., made of not tangible entities) or social (i.e. mix of physical and conceptual).

2.3 Taxonomies and technology trades-offs

Taxonomies are the key to decompose SoS in a manageable number of subsystems and components. This decomposition allows a direct comparison among different technologies suited for the same subsystems. Performing these comparisons is necessary to evaluate how different architectures perform in different missions, while looking also at how different infused technologies can influence mission success.

Historically, in the naval and aerospace fields, achieving this in a quantitative way has always been a challenge due to the vast number of components and connections (physical and virtual) present in vessels and aircrafts. In addition, by using hierarchical taxonomies, the problem of too many technologies was exacerbated. The lack of clarity in understanding deep components' relationships affected the satisfaction of basic missions' requirements. To limit this issue, in the past different navies have compared ships mostly qualitatively, stopping most of the quantitative analysis at a subsystem level. This lead to the lacking of studies on secondary effects of individual technologies present in the compared subsystems [54].

Observation 3: Historically, each subsystem was treated individually, reducing the intra-subsystems interactions captured and reducing the overall understanding of technology trades among architectures.

To solve the issue of having to deal with a very large number of technologies is possible by reducing the sample size to a manageable dimension by lumping together similar technologies to study their combined effects [55]. To study the effect, performance metrics are used as they allow the evaluation of different technologies – or pools of technologies –

in similar architectures. Since different technologies allow assets to perform different missions, by cascading quantitative parameters it is possible to evaluate mission performances. Historically this has been done only qualitatively, but with modern tools and higher computational power, a quantitative study is possible.

Observation 4: Technologies of the same family can be lumped together in one block to allow quantitative comparisons among ships with different subsystems and architectures.

2.4 Gaps in current taxonomies

Following Observations 1 and 2, it is clear there are some gaps in current practices that have to be addressed going forward. Independently from the solution chosen, it is important to reduce the compartmentalization among subsystems. This should allow a better understanding of the relationships among components, even when looking at different layers (physical, logical, etc.).

Using one taxonomy or another should not affect, in principle, the reality of the system. Nevertheless, it disturbs the way comparisons among subsystems are made. As discussed in Observation 3 and 4, historically technologies have been lumped together, according to different taxonomies, to perform trades-offs and comparisons among SoS. The need to lump together different technologies still stands. However, now this provides the opportunity to look at different ways of lumping those technologies together. In this sense, the role of taxonomies is even more important, as it affects how technologies are compared, and therefore how requirements for future investment are designed. In the context of this work, taxonomies are a fundamental piece as they drive how different fleets will be decomposed first, and compared later, to identify Robust Strategies.

Another important gap that emerges from previous observations is the need to ensure that taxonomies are comprehensive. In fact, many taxonomies seen in previous examples, focused only on specific areas and did not expand outside specific subsystems. While this gap could be interesting to cover, it is also a gap that can be easily addressed by setting clear problems' boundaries by decision makers. Once the boundaries of the problem are clearly set up, then it is just matter of mapping accuracy to ensure that all the components of the system-as-framed are included.

A more interesting challenge is to verify that the right inputs are collected when creating a specific taxonomy. In other words, how is it possible to ensure that all the relevant inputs are collected when a specific taxonomy (e.g., a physical taxonomy) is created. In this case, there are two possible approaches: the first one is to lay down the taxonomy and cross check it with another existing one to ensure that all aspects are covered, the second one is to rely on subject matter expert to ensure that all relevant pieces are there. The first approach is useful when the system is already known and when there are already verified taxonomies. A good example would be to use a functional taxonomy to ensure all functions are present when making a physical taxonomy. In the second approach, which was the one chosen to verify all the inputs for the use cases, as mentioned in the previous chapter, subject matter experts are consulted to ensure that all the expected relevant pieces are present. This doesn't mean that experts will have a say on how the taxonomy is used, but only that they ensure that all the relevant inputs are there to avoid the risk of having "garbage-in garbage-out" models.

2.5 Hybrid Taxonomies

When comprehensively describing large and complex SoS, it is not possible to focus on a single aspect. Complex systems have many internal and external interactions on multiple domains. As such, to capture those interactions it is necessary to look at many perspectives at the same time [56]. To fulfil this task one single taxonomy is not enough. Many taxonomies though, working together on different layers, can depict the full picture of a complex system. Different taxonomies work in combination as overlaying maps: the more there are in the stack, the more information are available. In trying to evaluate how different taxonomies working together support the deeper understanding of systems' interactions, the following question will be answered:

Research Question 1: Can new taxonomies of ships help in increasing the understanding of high-level interactions among components?

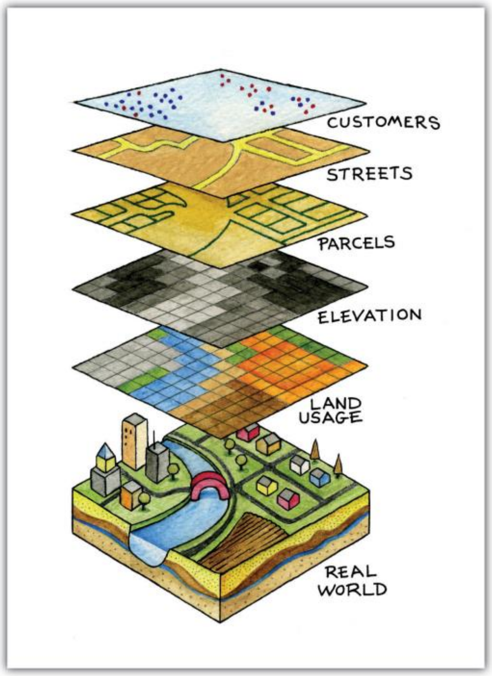


Figure 8: Overlaying taxonomies are like overlying maps. To have the Real- World view many are needed on the stack [57]

Looking back at the original naval problem it appears evident that multiple taxonomies are needed in order to coherently, and comprehensively, describe a fleet at sea.

Hypothesis 1: If a hybrid taxonomy, made by overlapping different taxonomies, is used to understand relationships inter-SoS, then it is possible to use it to comprehensively describe interactions among SoS.

In other words, Hypothesis 1 is saying that no taxonomy is superior to others when studying high level interactions, as for instance those among ships in fleet at sea. The only caveat is that the taxonomy used must be broad enough to account for the complexity of the system. To demonstrate this hypothesis a logical proof is provided.

2.5.1 Logical Case Study

If what described in Hypothesis 1 is not true, the conclusion would be that it is possible to holistically describe interactions inside complex systems (e.g., a ship) using one single type of taxonomy. In the following paragraphs it will be demonstrated how this is not possible due to the multi-domain nature of complex systems.

In formulating this demonstration, a missile defense system and its major subsystem will be used. Common missile defense systems are made of four major subsystems: the launching system, the missile, the command and control center and the detection system. The efficiency of the system is directly linked to the efficiency of each component of the kill chain. Therefore, investments looking at increasing the efficiency of the missile defense system as a whole will have to focus on each segment of the kill chain to achieve the desired effect.

When saying that the missile defense system can be the decomposed into four major subsystems an implicit taxonomy was taken – a functional one to be precise. The system was divided according to its four main functions: launching an effect, killing the target, controlling the sequence of operations, identifying enemies. When looking at this taxonomy it cannot be inferred how the physical components are divided, except the obvious fact that the missile must be independent from the control body. Therefore, from an investment perspective, one does not know if all the functions, except killing the target, are allocated in a single physical component or in many different components.

If a singular physical taxonomy is taken, and the same exercise of dividing the kill chain in many blocks is conducted, the results can be completely different from what was found in the functional taxonomy. It is not rare, in fact, that while the command and control center is integrated with at least one antenna for identifying targets, other antennas are present to illuminate the target from different positions. Following this logic, potential investments will now have to look at upgrading specific components without necessarily knowing the functional interactions with others. This means that investments will focus only on the selected hardware, without including the possibility of changing the functions performed by each component.

In summary, if only one taxonomy is considered, only a partial description of the complex system will be possible. This can cause the exclusion from the analysis on future needs of several critical aspects of the problem, and it can reduce the beneficial effect of the investments. Looking at the two taxonomies considered before, merging or distributing components and functions to increase the resilience of the system cannot be considered by any of the two taxonomies, as this will go outside of their boundaries. Therefore, it can be speculated that the solution achieved by any of the two taxonomies by themselves will be suboptimal to what could be achieved with a combined approach.

This case study showed how in a complex system several taxonomies must be taken into consideration at the same time, discarding the idea that one general taxonomy would provide sufficient understanding of the system. Therefore, this demonstration inherently supports the conclusion that multiple taxonomies are needed. While some may provide more critical information than others, they all complement each other by providing useful pieces of a complex puzzle.

CHAPTER 3. Modelling and Discovery Techniques for Naval Scenarios

In the past decades many efforts were made to achieve a better understanding of future naval scenarios. These approaches must be tied to how the US military currently prepares for future threats. As such, this chapter begins by briefly showcasing the current investment planning methodologies used by the US Armed Forces in order to answer Motivation Question 2. Using this knowledge, it is possible to understand which modelling technique is better suited to simulate the behavior and performances of future assets in forecasted scenarios. The chosen modelling technique will be applied to newly developed Scenario Discovery techniques, which will be able to quantitively find possible Vulnerable Scenario. In the last part of the chapter a technique called Principal Component Analysis will be introduced. This will be used to augment the Scenario Discovery algorithm in complex cases.

Motivation Question 2: What are current investment decision making methodologies used by the DoD?

3.1 Current investment planning methodology in the US

To achieve a defense investment strategy, different documents, such as the National Security Strategy, the National Defense Strategy, the National Military Strategy and the Defense Planning Strategy must be translated into requirements, desired capabilities and foreseen tactics for a comprehensive force structure.

DOD planning efforts are built around a doctrine called *Predict-then-Act* [58]. In this doctrine analysts are required to plan for the future, including strong assumptions regarding how the world will evolve, affecting the analytical rigorousness of the planning effort. Therefore, whatever solution is obtained, the outcome should not be seen as an absolute truth, but rather looked at with critical eyes.

According to Mazarr [47], current defense planning is mostly demand-based. Meaning that requirements and constraints of envisioned future engagements drive the development of strategies, capabilities, and capacities.

Mazarr assesses those two main types of planning are present in the current DoD doctrine: Demand-Based Planning and Supply-Based Planning. The first one is the most prominent and represents what many have criticized as a “Cold War approach” [59], as it requires a clear understanding of who the enemy is, and what capabilities and objective(s) does it have. The other approach – Supply-Based Planning – is more focused on specific real-world restrictions like current fleet size, budget limitations and capability mix. Mazarr supports the theory that in order to achieve a balanced force, it is advantageous to use both types of planning methods during different phases of the study, as they provide complementary insights to the process.

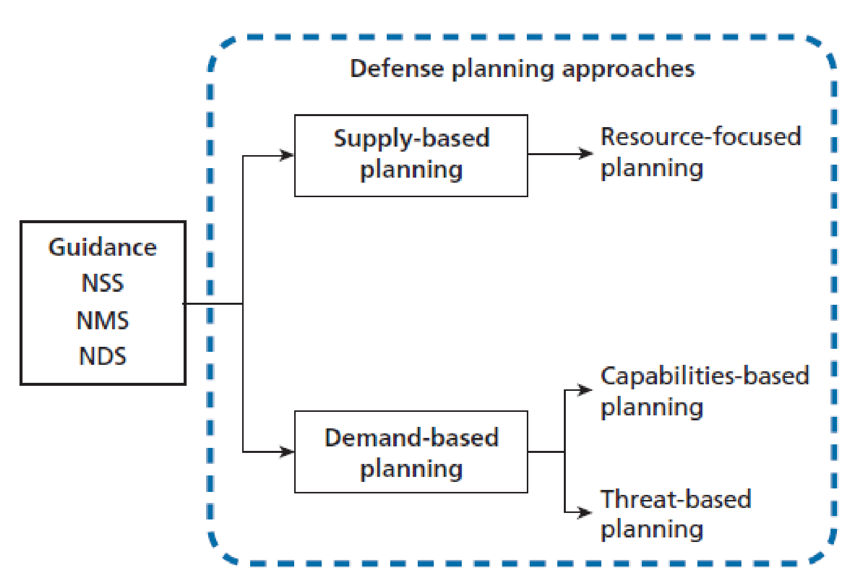


Figure 9 Defense Planning Approaches [47]

Mazarr concludes his report advocating that the distance between the two planning types is only part of the problem, and that the DoD doctrine used in the scenario development process creates a gap. The process is in fact inflexible, costly and time consuming to the point that policymakers are forced to use heavy sets of assumptions to stay within their assigned boundaries.

Observation 5: Traditional methods lead toward a lack of characterization of all the different dimensions of uncertainty by introducing several assumptions in the planning effort.

3.1.1 Demand-Based Planning

Demand-Based Planning has been the chosen planning method by the DoD in the past years. In general, Demand-Based Planning targets decision makers' objectives on potential future conflicts and threats. To shape Demand-Based approaches and to size its armed forces, the US has been following a “two-war” approach since the 1960s [47]. This approach, originally drafted during the Kennedy administration, has largely evolved from the original objective of being able to fight two major wars and one limited conflict, to include specific threats, such as terrorism. The “two-wars” approach also drives some of the required missions the US Armed Forces [60] should be able to perform:

- Defend the Homeland
- Perform sustained and distributed operations against terroristic groups
- Deter aggressions and ensure alliances via a forward-oriented posture

In case of failure of deterrence, the role of the US Armed Forces shifts toward a more active one, in which the military should be able to engage and defeat a regional power,[1] while denying the objectives of another adversary in a different theatre.

Due to their broad scope, Demand-Based approaches fit well in the general, larger national defense planning, providing high-level design requirements and high-level strategic views in the early planning process. Nevertheless, focusing too much on the broad-level picture can lead to producing impractical results in the end. Moreover, if the objective of the national strategy is to counter specific enemies, the results produced will be tailored against specific threats leaving gaps open in other theatres.

[1] Note that the US currently recognize only itself as a global power, therefore adversaries like China or Russia are considered either regional or emerging powers.

Observation 6: Investment solutions focused on a narrow environment increase the risk for single point solutions.

3.1.1.1 Threat-Based Planning

Threat-Based approaches are designed to respond to specific, potential, future enemies and conflicts. Force structures are therefore planned and shaped against specific targets, while high-level strategy documents provided the basis for these models. Threat-Based planning was common in the Cold War era when the world was divided in two clearly marked spheres of influence [61]. Today this is not true. State and non-state actors use more gray area tactics across all platforms and domains to cloud where the real threat is.

In general Threat-Based approaches encourage a more in-depth study of enemies' capabilities. If a conflict arises, thanks to this approach, resources can be focused on what is critically needed against specific threats. Moreover, if scenarios are modelled, planners can include uncertainties and unique societal and political aspects present in a real-world conflict. Unfortunately, all the technical and planning information come from intelligence gathering. This means that information is inherently ambiguous, and could be unreliable, or even purposely leaked to disrupt planning.

This approach is information dependent; therefore, it can lead to a reactive culture in which single point solution responses are created losing sight of the global picture and of other potential threats.

Observation 7: Threat-Based planning does not capture all the different dimensions of uncertainty present in scenarios. As such, there is a high risk of generating single point solutions.

Observation 8: Investment solutions generated relying solely on threats – perceived or actual – can lead to self-fulfilling requirements not reflecting the real threat environment.

3.1.1.2 Capability-Based Planning

Capability-Based Planning approaches focus on identifying general capabilities that will be needed in a less-defined war scenario rather than finding specific enemies and threats [62]. These approaches are therefore looking at those technologies that can meet high level sets of objectives based on national strategies. This allows for well-structured investment plans in which the system is decomposed on different operational levels. Goals are often divided in time and full capability is achieved step by step.

The main advantage of these approaches is that it does not require planners to agree on hypotheses regarding the set of potential adversaries. Reducing the risk of delivering products overly focused on a single problem and shifting the investment plan from being responsive to being prepositive. However, too much abstraction can increase the risk of a fit-for-all solution which leads to requirements growth. It can also drive the removal of the adversary political and societal connotations from the planning process, pushing it toward one that focuses only on the CONOPS. Looking only at CONOPS results in the lack of important contextual facts, leading to plans that are impractical, unfeasible, and generally not credible against real-world adversaries.

Observation 9: Capability-Based Planning can lead to investment programs with excessively abstracted threats that have impractical, not credible and unfeasible solutions.

3.1.2 Supply-Based Planning

Differently from Demand-Based Planning, Supply-Based plans are bottom-up approaches focused on today's status of the world. Usually, information to start the analysis includes current force size, budgetary limitations, and capability mixes available to build up forces. The most important of these approaches is Resource-Focused Planning.

Resource-Focused Planning uses today's constraints to identify possible investment patterns for the future. The status quo is used as a baseline and desired capabilities are built on it limited on what is available today for investments. For this reason, this approach is often called budget-driven planning, as the goal is to use the budget as a limitation to prioritize certain capabilities as desired by the stakeholders.

The reason why this bottom-up method is commonly used is to root high-level strategies to resources available for investments, reducing the risk of unfeasible solutions. In Resource-Based approaches, the focus is to modify already existing defense programs in marginal, but meaningful, ways. Being so close to already existing assets, designed improvement plans are easy to translate into applicable ones. Unfortunately, for the same reasons, these approaches might fail to adapt to future issues leading to more capability gaps.

Like Threat-Based planning, these approaches are focused on something that already exists, therefore the solution could be too tailored for a single theater, rather than targeting a long-term global problem. With that, Clark et al. have said that even if Resource-Focused planning is imperfect, it reflects how planning is conducted today to a large degree [24].

Observation 10: Bottom-up methods as Resource-Focused ones are heavily dedicated to the specific needs of today, losing sight of the global scenario and of long-term objectives.

Observation 11: Bottom-up methods as Resource-Focused ones are constrained in technology improvements by what is available today, limiting therefore the possible response in covering technological gaps on the long-term scenario.

3.2 Modelling techniques for quantitative naval scenarios

Being able to see how different approaches might impact future engagements its key in deciding on what to invest in. As such, in the past years many started using scenarios. Scenarios are among the most useful tools in defense planning, so they are largely used today by DoD [63], mostly in Demand-Based planning. In this context, scenarios are defined as a set of assumptions and information that are used to simulate the problem at stake.

Scenarios can be used to link threats and capabilities to assess them and make tradeoffs. They allow for more reality-based analysis in which planners can look at factors like the size and type of threats, the operational terrain, and engagement concepts. Decision makers can test different CONOPS by changing warning times, posture and the deployability of assets [61].

As a critical part of planning, scenarios must be carefully designed to avoid skewing results in one direction or another. Statistical analysis tools and agreed-on-results methodologies can mitigate these risks [64] by providing important information on the scenario. These tools can be vital in extracting information from clouded data allowing decision makers to be able to visualize what are the macro effects happening in the scenario. Moreover, by employing agreed-on-results methodologies, decision makers are less influenced by assumptions that might be not verified. To be able to reproduce these scenarios, different modelling options are available, each tailored to highlight specific aspects of the modeled scenario, but as mentioned in Research Question 2 not all are best suited for naval scenarios.

Research Question 2: What modelling techniques can be used to quantitively simulate naval future scenarios?

3.2.1 System Dynamics

System Dynamics modelling [72][73] is a common modelling technique specifically designed for strong interactions between different parties in the model. The focus of this modelling technique is on the whole scenario rather than individual pieces. Because of its focus on scenarios, many have said that System Dynamics can be matched very well with DoDAF. Nevertheless, since it lacks multi-level fidelity – the capability of looking at several assets while looking at the scenario – it does not allow for the monitoring of specific assets but only for the processes happening.

3.2.2 Discrete Events

Discrete Event modelling [67] is useful when use cases have sequential tasks and events. In its classical formulation, different assets, or entities, are passive as they only manage those attributes that affect them. Entities retain the ability to change the way they handle information. This change can happen as often as every time step. The reason why Discrete Event modelling was discarded in the end discarded is that it was generally more focused on the process rather than on the assets, and so is not best suited for looking at technology investments.

3.2.3 Agent-Based Modelling

The modelling method selected for this thesis is Agent Based Modelling [68]. This modelling technique involves agents which can act autonomously within the simulation. Each agent has its own goals, capabilities, and behaviors. They are part of the fleet, but also stand-alone assets. The lack of a general central authority allows in fact for a better representation of systems interactions, in which complex dynamics and emergent behavior can be captured and analyzed. While this modelling technique is focused on agents, it does not lose track of the bigger picture. In fact, it allows the modeler to look at each agent individually and at the scenario at the same time. The main disadvantage though, is that in very complex missions the computational cost might be hard to reduce without losing relevant information.

Agents, being autonomous and independent, lacking a central authority but yet being part of a group and working together, are a great representation of ships at sea in a fleet. As previously mentioned at the beginning of this paragraph, the chosen modelling method is Agent-Based Modelling. This is now clearly formulated in the following assumption:

Assumption 1: Agent-Based Modelling is the most suited technique to model the independent behavior of assets within scenarios, when looking for possible technological investments.

3.3 Scenario Discovery techniques

Modelling and simulations are needed to compute the interactions happening inside scenarios to understand the role of each technology of interest. But, to find critical variables, and to identify possible dangerous futures more analyses are needed. But what technique should be used to find the vulnerable scenarios when dealing with large se of uncertainties?

Research Question 3: What technique should be used to find vulnerable scenarios in a large dataset with deep uncertainty on the future evolution?

While different methods are available, due to the deep uncertainty factor it was conjectured that Scenario Discovery is the chosen group of techniques that is used to perform this task.

Conjecture 1: Scenario Discovery techniques will be used to find Vulnerable Scenarios.

As described by R. Lampert on the RAND Future Methodology webpage[2]: *"Scenario Discovery uses statistical/data-mining algorithms to find policy-relevant clusters of cases in large, multi-dimensional databases of simulation model results. Conveniently*

[2] https://www.rand.org/pardee/pubs/futures_method/scenario_discovery.html

interpreted as scenarios, these clusters help illuminate and quantify the tradeoffs among alternative strategies under deep uncertainty."

Across the field two main techniques have emerged: Classification and Regression Tree (CART) [69] and the Patient Rule Induction Method (PRIM) [70]. The are many advantages to using Scenario Discovery methods, and as such, the planning communities are now shifting toward them. Scenarios Discovery provides a way to tie in a more realistic way of investing to the real world, allowing a more in-depth participation of stakeholders. This enables a better prioritization of needed technologies and capabilities. Scenario Discovery also helps in providing a time element, thus pushing planners to look at parameters during peak response times and during sustained operations, providing a better understanding of CONOPS, force readiness and sustained operations. Being part of the Robust Design Methodology family, Scenario Discovery methods are designed to be "agree-on-results" method, thus they reduce the effect of hypothesis in scenarios [58] helping decision makers in taking less biased decisions. In helping decision makers to find the most Vulnerable Scenarios, the first choice is the one regarding which of the two Scenario Discovery method to use.

Research Question 4: What Scenario Discovery method can be used to find Vulnerable Scenarios for naval fleets in a credible and rapid way?

Between the two main techniques available, PRIM and CART, it was decided to use PRIM for this thesis. In fact, while some say that CART achieves better performances in identifying subgroups of interests in large datasets [71], even though this is not always the case as proven by Kwakkel [72], those studies rely on the hypothesis that a precise subgroup is desired. On the contrary, in this thesis the interest is to screen large portion of the design space to identify relevant policies. There is no interest in being extremely precise with the cases identified, especially given the fact that an iterative algorithm will later on modify and upgrade what was originally found. The other identified difference is the "greediness" of the algorithm when peeling data. As it will be shown in the next section, PRIM controls the thickness of the layer peeled at every iteration and therefore it can advance rapidly if few cases of interest are present. On the contrary, CART follows an optimization criterion for splitting which doesn't account for minim steps, leading to slower iterations. Therefore, considering also speed as one of the critical factors to take into account, PRIM was chosen. Patient Rule Induction Method (PRIM)

PRIM was introduced in 1999 by Friedman and Fisher [73] as a data mining technique with the objective of finding regions in the design space with extreme values of performance metrics. In this sense, this method creates rectangular regions by using simple rules on the performance metrics defining the design space. The idea behind PRIM is to iteratively find smaller and smaller sets of the sample space that can construct an optimal rectangle around the solutions not satisfying a decision criterion. In finding this "optimal" rectangle the algorithm is trying to keep the most relevant characteristics of the sample set to maintain results consistency. This is done to safeguard against taking hasty decisions. In fact, all suboptimal rectangles are stored to allow decision makers to go back to a previous step. PRIM works with both continuous and discrete variables but depending on the sample size and on the amounts of variables. Three techniques are possible for creating the scenario boxes: peeling, pasting or covering [70].

Peeling works by removing a strip of the design space that does not contain strategies that have failed the scenario in each iteration. The first box is the full scenario, this is reduced by a certain quantile every iteration depending on the number of failing strategies present in the box. To do so a conditional distribution F(·|S) is used – where S is the subset of the whole scenario box, therefore, when removing the k-th candidate box b this needs to satisfy the following:

$$F(b|S) = \alpha\,(1 - \alpha)\,k - 1$$

The peeling iterations are done until the box B satisfies F(B|S) ≥ β. The parameters α and β represent respectively the step peeled at each iteration and a tuning parameter for the size of the box.

Pasting can be used to adjust the outcome of the peeling effort. This concept is the inverse of peeling: starting from the result box B layers are pasted along the boundary increasing the candidate box side b. This happens as long as the average of non-successful strategies increases. This method has the disadvantage of not being able to control the result box, leading to relatively large sample space. As such some have proposed an application that substitute the peeling + pasting approach with jittering. Jittering adds and subtracts at the same time layers to the original scenario box in such a way to keep increasing the average of the box. This reduces the risk of cancelling out information around local minimums by adding parts of the design space that carry significant information.

Covering is similar to the peeling + pasting approach and it leads to a final region which is the union of the two boxes obtained by peeling and then pasting layers. The difference is that instead of doing all the peeling first and all the pasting second, in covering peeling and pasting are executed in sequence at every layer, each time removing therefore the optimal outcome of the previous iteration.

Assumption 2: Given the sample size and the number of variables, Covering is the PRIM technique that is used in this study

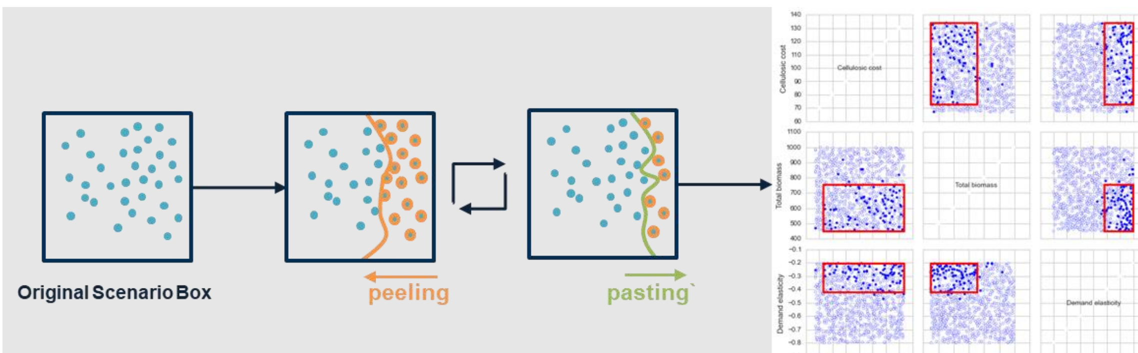


Figure 10: Covering is the result of sequential peeling and pasting operations over a scenario box, in red on the last picture on the right a set of PRIM scenario boxes

PRIM works by evaluating a binary value from the output of the simulation. The user needs to define a threshold value to identify *cases of interest* from less interesting ones. These cases could be those in which a fleet is successful in achieving a task. PRIM uses this binary condition to filter the output database. In this process different PRIM box sets are generated, each of this box is displayed as a point in a trade-off graph generated as output of the PRIM algorithm. The user can select the box set that is more appropriate and useful for the application. Each box set contains several PRIM boxes, each PRIM box contains a certain number of *cases of interest*, and the related performance metrics value ranges.

In the trade-off diagram previously mentioned there are 3 main parameters: Coverage, Density and Restricted Dimensions. Coverage is the ratio of *cases of interest* (i.e., when the fleet succeeds or fails) inside a box set to the total number of *cases of interest*. Coverage is important as a high coverage means that the found scenarios can explain more *cases of interest*. Density is the ratio of cases *of interest* inside a box set to the number of *cases of interest* within that box set. Density is important because a high-density means being able to use each scenario as a strong predictor for other *cases of interest*. Ideally one would like to have both high density and high coverage, unfortunately though these parameters are usually in tension and increasing one often reduces the other. Restricted Dimensions are the number of variables that are sequentially restricted during peeling operations. Usually, the higher this number the lower the coverage is – as more and more *cases of interests* are excluded from the box set.

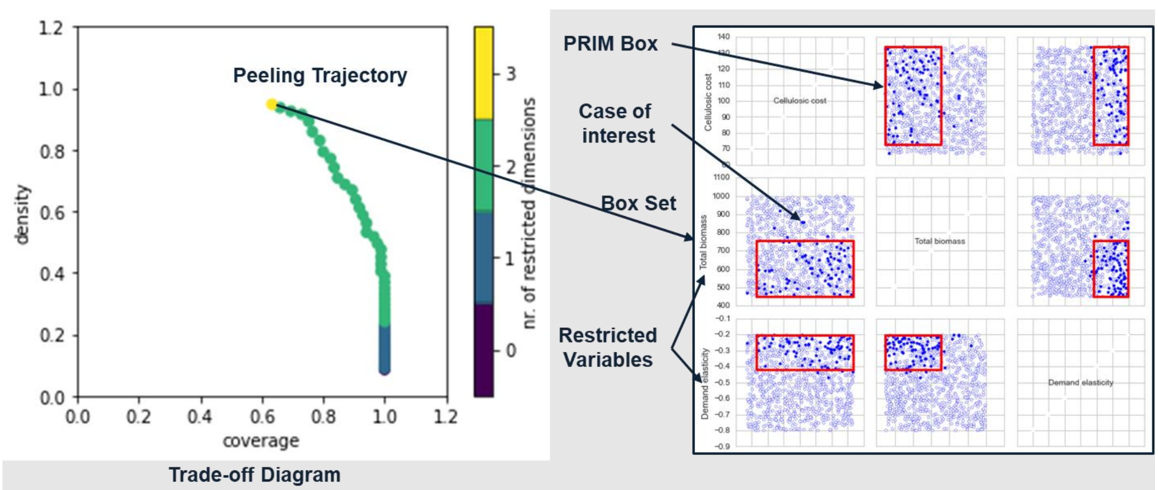


Figure 11: PRIM nomenclature. The peeling trajectory is displayed in the trade-off diagram. Each point in the peeling trajectory is a box set the user can select. Each box

set has its own density, coverage, and restricted dimensions. Each box set has many PRIM boxes inside. In dark blue, there are *cases of interest*.

3.4 Principal Component Analysis (PCA)

PCA can be used with PRIM as a preprocessing step. Often, the data used in PRIM are scattered and with design spaces hardly hyper-rectangular. As such, the Scenario Discovery method can have issues converging and finding optima hyper-rectangular regions. PCA offers a convenient way to rotate the reference system of the data providing a better framework for PRIM to operate [74].

There are at least 2 main versions of PCA that can be employed with PRIM. In the first one, PCA-PRIM, all the uncertainty axes are allowed to rotate. In the second one CPCA-PRIM, where the C is for constrained PCA, only user-domain parameters are allowed to rotate. This second method produces clearer results for the user as the modified parameters are only the one that the user uses. In contrast, in PCA-PRIM even dissimilar parameters and their combination can rotate, leading to much more complex results [75].

In general, both follow the same approach [74]. First, the method is applied only to the cases that failed to find the new set of rotational coordinates. Next, the full dataset is rotated using the discovered coordinates. This process is basically taking the covariance matrix and it is rotating its eigenspace. By doing so, the new axes of the full dataset are better aligned with those of the failed cases subset. This helps PRIM in finding better box sets with higher coverage and density. The measure of improvement is the harmonic mean of coverage and density [76].

It is worth mentioning that since PCA is a linear technique in some cases it will not produces results acceptable for PRIM. While none of those cases was encountered in this work, there are also other quadratic methods that could be used when PCA fails to produce good results. These techniques include Quadratic Regularized PCA and Quadratic Discriminant Analysis.

3.5 Gaps and Conclusions

In this chapter the focus was on the different methods to model and simulate military operations, and on how that information can be used to discover future high-risk scenarios for a specific fleet of interest.

Looking at currently investment planning practices it emerged that there is a lack of characterization for all the different elements of uncertainty in the scenarios. This leads to a need for better understanding what are those that in the future will be Vulnerable Scenarios. Answering this point in a *quantitative* way will be the key to find investment solutions that will be able to tackle the reality of the future.

To do so, the chapter focused on understanding what the methodologies available today can do, and which one is the most appropriate one for this problem. After looking across several modelling techniques Agent-Based modelling was selected. ABM has several advantages compared to others, and it also allows the formulation of low fidelity model – increasing the speed of the modelling effort while retaining quality in the outcomes.

Having a modelling technique is not enough though. The output of that modelling and simulation effort must be analyzed to find which factor will be the most critical one in the future. To complete this task Scenario Discovery methods were considered, with a particular focus on PRIM.

SD and PRIM have often been used to evaluate different policies, but they have never been used to evaluate quantitative naval scenarios. This is one of the main gaps that was discovered throughout the literature review effort. Looking at the potential of these techniques and knowing that it has been used successfully in other military technology investment planning efforts (e.g., ammunition), it is believed that PRIM, if used together with Agent-Based modelling can find vulnerable naval scenarios. Moreover, it is believed that this can be done in a quantitative, rapid, and credible way. The demonstration of this hypothesis will be conducted later with an ad-hoc experiment. From what described so far on the roles of PRIM and Agent-Based modelling, and their ability to identify naval vulnerable scenarios, hypothesis 2 follows:

Hypothesis 2: If PRIM is used together with Agent-Based modelling, then it can rapidly and credibly identify quantitative sets of naval Vulnerable Scenario

CHAPTER 4. Algorithms for Technology Selection

In Chapter 2, different taxonomies were described to show how the same system, or system of system, can be decomposed and what are the consequences of picking one decomposition over another one. In Chapter 3 different acquisition methodologies were discussed, the focus being the use of scenarios and how they shape future assets acquisition.

In this chapter we will rely on what was discussed before and focus back on the technology side of the problem. The effort is to answer the second part of the research objective of looking at “trade-offs among naval assets and technology”. To do so, the first step is to understand how assets are currently quantitatively compared in investment strategies. Motivation Question 3 gives a good start to see which gaps are present in this field, and where they are.

Motivation Question 3: How are assets currently quantitively compared in investment methodologies?

4.1 Technology Comparison

In Chapter 2 by studying different taxonomies, the goal was to understand possible assets’ decompositions better. This now enables us to look for only specific technologies whose benefits cascade to a significant number of subsystems. By understanding which taxonomical combination provides the most benefits to study, the goal shifts toward understanding how to perform technology trades in a quantitative way.

Historically, the extreme complexity of naval systems, the differences in design practices and the development of many ad-hoc variants have been a showstopper for quantitative comparisons and trade-offs analysis [77]. Moreover, the presence of a multitude of different components and non-comparable subsystems did not help downscaling design spaces to manageable ones. It is therefore of great interest, the ability to pinpoint at the benefits that different technologies provide and at the gaps they cover in a quantitative way.

Ships comparisons were often done among allied navies to see how tactical choices were reflected on the design to facilitate technology transfers [55]. These comparisons were carried out at different level by looking at how vessels of the same class differ in various subsystems [54]. Each subsystem was treated independently, and the connections between technologies onboard and the ship's mission requirements were drafted. From these practices, the critical role of taxonomies in trade-offs emerges even more.

In traditional comparisons, each ship is decomposed in different subsystems that are again decomposed in basic technology bricks. These are later lumped together allowing comparison of inherently different architectures from both the technology point of view[54] and the crew sizing one [78]. Each technology brick, or pool of technologies, directly contributes to a specific task as shown in the taxonomies discussed in Chapter 2. By comparing high level metrics, such as ranges or VLS cells, ships can be compared, and subsystems and technologies traded in a quantitative fashion. Each exchanged piece of technology provides different performance metrics, and its characteristics are reflected on the tasks the ship can perform.

All of this leads to two observations:

Observation 12: Historically each subsystem was often treated individually, reducing the amount of captured intra-subsystem interactions.

Observation 13: Technologies of the same family can be lumped together in one block to allow quantitative comparisons among ships with different subsystems and architectures.

From these 2 observations and from the literature review conducted one gap emerges: the need to be able to compare technologies, within assets, to be able to compare assets' performances. Moreover, to evolve past the historical approach this comparison must be done in a practical, credible, and quantitative way. The gap is summarized as follows:

Gap 1: From historical trade-offs in the naval field the need to compare technologies, among assets arises. This comparison must be conducted in a quantitative, practical, credible and way to understand the impact of each technology on the overall ship, or fleet.

4.2 Technology Pools Identification

When looking at Gap 1, the first question arising is how to select the first pool of technologies. To select the appropriate technology bundle, a statistical relevance and testing evaluation must be used to enhance the long-term robustness of the fleet. Which provides the grounds for Research Question 1, as follows:

Research Question 5: What tool can be used to select sets of naval technologies to enhance fleets' long-term robustness?

Looking at the previous section, it emerges that whatever solution chosen must fulfil some requirements to advance the state-of-the-art. First, the tool chosen must address technologies in a quantitative way. This means that each technology will be decomposed in a series of parameters that can be quantitatively used by the model in the simulation. The direct consequence of this is that those parameters will be used as proxies for the different technologies of interest. Ideally, by changing the value of one parameter it could be possible to go from one technology to another. For example, if the parameter of interest is shooting range, a shooting range of 10 km would imply using a short-range missile, while a value of 200 km would imply a long-range missile. This ensures that whatever method is chosen it is technology agnostic.

A second requirement is adaptability. There is little interest in selecting a static pool of technologies; on the contrary, it is much more interesting to be able to modify that pool so that the fleet can advance even in losing scenarios. This means that the initial set of values for each technology should not be a fixed constrained, but rather a starting point.

The third requirement is the speed of the whole process. This work is by no means trying to use high fidelity modelling – which will increase computational efforts and time. On the contrary, one of the main characteristics is the interest in using a low fidelity approach to gather insightful information to help reducing the size of the design space for high fidelity modelers, which will anyway later verify whatever technology is selected. As such, the method that is chosen should be able to analyze large design spaces with a quick turnaround. This requirement is in line with efforts of the US Air Force to contain the turnaround time for quick simulation evaluations to 90 days [42].

Following all these requirements it emerged that a hybrid approach, rather than a single technique was the right tool. A large experimental simulation will be used as a starting point, but differently from conventional uses of Design of Experiment (DoE) it will employ an iterative algorithm to adapt the DoE to the evolving scenario. This way it is possible to look at the broad spectrum of the technologies of interest, while being able to adapt them to the scenarios in an iterative way.

Assumption 3: Given the size and type of problem a hybrid approach employing both an iterative algorithm and a large DoE will be used

4.2.1 Hybrid Approach

To answer all the 3 requirements established, a DoE is not sufficient due to its static nature. In fact, solely relying on a DoE does not allow the user to expand and evolve considering the history of the metrics of interest, but it only allows the user to have a wide and static starting point. Therefore, the preferred option fell on a hybrid approach in which the DoE is complemented with an iterative algorithm which will update it to account for new information from the simulation itself.

The starting point of this approach, as mentioned, should be a DoE. Employing a DoE helps in moving away from traditional database and historical regression models which are no longer valid [72] [73]. This is important in this study as the focus is the research of technologies which might fall outside the boundary of already existing models, leading to unacceptable answers. The DoE will contain values for specific technological parameters, ensuring that the effort remains quantitative. Once the DoE is available the information will be plugged in the model for the simulation part.

At this stage, the second requirement must be satisfied. Whatever modelling and simulation tool is used it must be tied with an algorithm that allows the model to learn simulation after simulation. This algorithm should connect the output with the input in such a fashion that iteration after iteration the fleet evolves, and it is eventually able to overcome the task.

Because there is an interest in moving away from known technological constraints, this algorithm should also ignore the original design space limits imposed in the starting DoE. In other words, it should let the fleet evolve freely (while still focusing on using the technologies of interest). The decision maker in the end will evaluate the resulting fleets to pick whatever meets his criteria.

One possibility would be to do this using a much larger DoE and then test all the cases without the need to use an iterative algorithm. If this approach is pursued, to include all the cases that the iterative algorithm would be able to test, the DoE's boundaries will have to be expanded a lot more than the original design limit. This will increase computational time significantly, especially in those situations with many variables. Moreover, since the evolution of the parameters is unknown it might still require going outside of the newly defined DoE boundaries.

4.3 Hybrid approaches to technology selection

Following A second requirement is adaptability. There is little interest in selecting a static pool of technologies; on the contrary, it is much more interesting to be able to modify that pool so that the fleet can advance even in losing scenarios. This means that the initial set of values for each technology should not be a fixed constrained, but rather a starting point.

The third requirement is the speed of the whole process. This work is by no means trying to use high fidelity modelling – which will increase computational efforts and time. On the contrary, one of the main characteristics is the interest in using a low fidelity approach to gather insightful information to help reducing the size of the design space for high fidelity modelers, which will anyway later verify whatever technology is selected. As such, the method that is chosen should be able to analyze large design spaces with a quick turnaround. This requirement is in line with efforts of the US Air Force to contain the turnaround time for quick simulation evaluations to 90 days [42].

Following all these requirements it emerged that a hybrid approach, rather than a single technique was the right tool. A large experimental simulation will be used as a starting point, but differently from conventional uses of Design of Experiment (DoE) it will employ an iterative algorithm to adapt the DoE to the evolving scenario. This way it is possible to look at the broad spectrum of the technologies of interest, while being able to adapt them to the scenarios in an iterative way.

Assumption 3 and Gap 1, the focus now shifts in identifying which type of algorithm can cover the aforementioned gap. This decision should not be taken just in the context of finding and selecting technologies, but it should also consider the whole idea of this thesis – i.e., finding naval technology to support R&T in future Vulnerable Scenario. Although the choice and requirements of selecting the type of DoE will be discussed later, the large

DoE hybrid approach leads to the next research question to determine an integrable algorithm

Research Question 6: Which hybrid approach can be used for quantitatively selecting technologies to invest in a naval fleet in a credible, practical, and rapid way?

Lempert has had a prolific research activity in this field, so to answer the previous question the effort started with understanding his work: Robust Decision Making (RDM) [81]. The other two methodologies that were analyzed are also part of the same family, but they focus on adaptive policies [82] and multi objective analysis [83] respectively. The reason why across all the options the analysis focused on RDM-like methodologies is that these can deal with large set of uncertainties. All three techniques in fact allow the decision maker to work without really knowing all the assumptions of both scenarios and technologies of interest. This enables the decision maker to explore different scenarios and to investigate how different technologies will perform in an unbiased way.

4.3.1 Robust Decision Making

Robust Decision Making is the first technique of the family. Throughout the years, there have been multiple spinoffs of this method, tailored to solve specific issues. The core of RDM is the idea that thousands of experiments can be run to calculate how different uncertainties affect the results [84]. These experiments are performed by changing the variables of interest each time. Variables can be either at the scenario level or at technology level. By doing this RDM can look at the problem two ways: what happens if the scenario studied is not exactly what was expected, and what happens if the technologies of interest did not perform as planned.

Lempert reports that all RDM techniques have at least four basic steps [81] as reported in Figure 12. Decision Structuring is the phase in which all needed information is gathered. In Case Generation the data from phase 1 are used to generate the scenarios and the pool of strategies to be later analyzed; a DoE can be used as substitute of this phase. Phase 3 is focused on Scenario Discovery, which was explained in detail in Chapter 3. Finally, in phase 4 the information on the scenarios is combined with those on the technologies to find which strategy works the best across more scenarios.

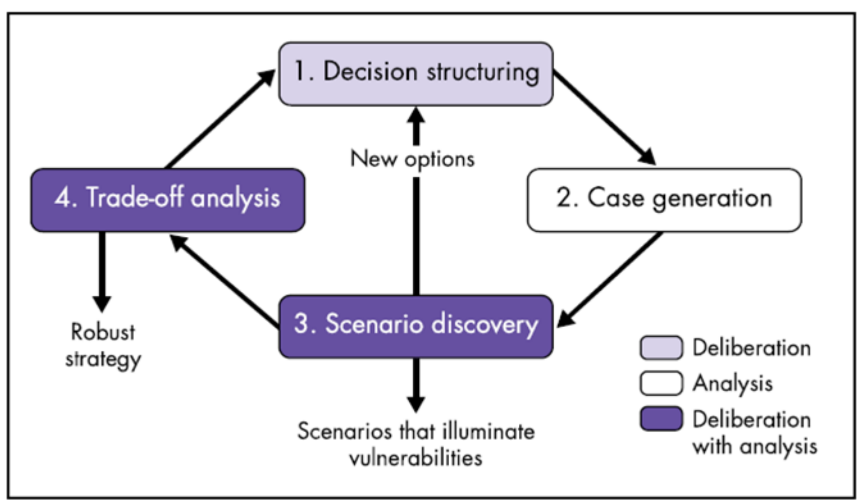


Figure 12: The four steps of RDM: Decision Structuring, Case Generation, Scenario Discovery and Trade-off Analysis [81]**.**

RDM is considered an "agree-on-decision" method, which means that all the strategies, scenarios and technologies are considered before analyzing the results. This is in contrast with more often-employed "agree-on-assumptions" methods, in which the design space is reduced by setting tight ranges on assumptions. In RDM after all the strategies are considered, models are run to test the different options in a large manner. This way more different futures and technological options (also called strategies) than “agree-on-assumption” techniques can be found.

The goals of RDM can be summarized in two main points: first, the method is looking for robust and not at optimal strategies [85]. This means that the effort is placed on looking for strategies that will work on a broad set of possible futures. Secondly, to achieve robustness strategies should be at least partially adaptive. This means that they should be able to evolve overtime to account for new information [85], either inputted by the user or generated by the model.

Observation 14: In RDM, the focus is more on predictive failure rather than strategies that can support success.

4.3.2 Adaptive Robust Decision Making

In dealing with scenarios surrounded by deep uncertainty, Hamarat has developed Adaptive Robust Decision Making (ARDM) [82]. This technique, as the previous one, is an iterative model-based approach to evaluate policy or technology strategies in scenarios with deep uncertainty on the variables. To be more explicit, deep uncertainty happens when the evolution of future scenarios is not clear, nor can the models properly represent this evolution [85].

ARDM uses a simple but powerful concept: *signposts and triggers*. Signposts are used to track selected variables while the scenarios evolve, each metric of interest has its own signpost. If the tracked variable reaches a certain threshold the trigger is activated. Triggers are actions that influence the inputs of the next simulation. Four actions were defined by Hamarat [82]: Reassessment is the action in which the original strategy's assumptions are checked for validity; Corrective is the action where adjustments to the basic strategies are made; in Defensive actions the basic strategy is reinforced to preserve its benefits; and finally the Capitalizing action aims at reducing the weight on the basic strategy to exploit opportunities in the design space. The pursued action is reflected on the input of the next simulation. Through this recursive approach, the model can evolve simulation after simulation, providing as output strategies that have adapted to scenarios proving to be robust and mature.

Different from pure RDM, with ARDM decision makers can find both promising and troublesome regions. But doing this in a quantitative way requires very high computational power. In fact, in many of the studies conducted using this method either the number of variables was low, or the model was simplistic [86].

Observation 15: ARDM requires a lot of computational power to be implemented, and it lacks the ability to handle more than a few variables.

4.3.3 Multi-Objective Robust Decision Making

Multi-Objective Robust Decision Making (MORDM) [83] is a spinoff of RDM developed by Kasprzyk et al. that includes a multi-objective optimization step. This step is conducted at the beginning of the process, and it is performed to look for solutions before starting the exploratory modeling and the scenario discovery phase [87].

One of the differences of MORDM from the other RDM methods is the use of approximate Pareto frontiers. MORDM, to avoid narrowing the set of objectives creates a buffer zone around each Pareto point, meaning that the optimal solution is not needed nor desired. This helps a lot in reducing computational efforts in a methodology that is not looking at one optimal solution, but rather at a solution that is robust across several scenarios.

Differently from the other RDM methodologies, MORDM can perform well with a high number of variables. This is done by employing Multi-Objective Evolutionary Algorithms in phase 2 of the RDM general method presented in Figure 12. By doing so the next step receives a rich set of alternatives, which do not require assumptions for the decision maker to start the trade-off analysis [83].

The main issue with this technique though is that it struggles with adaptiveness. In fact, while robust strategies can still be found, due to the high number of variables it’s hard for the algorithm to account for incoming information. Moreover, some have criticized this method saying that it leads to overly complex models in which there are so many uncertainties that it is hard to frame which one is relevant and which one is not [88].

Observation 16: MORDM is not an adaptive technique, and it struggles to frame problems with large uncertatintes.

4.4 A New Hybrid Approach

This chapter started with work on how technologies are compared, historically and today, to understand how predictions on fleets’ long–term robustness are made. The following step was to find ways to select those technologies and to test them in a modelling environment. In A second requirement is adaptability. There is little interest in selecting a static pool of technologies; on the contrary, it is much more interesting to be able to modify that pool so that the fleet can advance even in losing scenarios. This means that the initial set of values for each technology should not be a fixed constrained, but rather a starting point.

The third requirement is the speed of the whole process. This work is by no means trying to use high fidelity modelling – which will increase computational efforts and time. On the contrary, one of the main characteristics is the interest in using a low fidelity approach to gather insightful information to help reducing the size of the design space for high fidelity modelers, which will anyway later verify whatever technology is selected. As such, the method that is chosen should be able to analyze large design spaces with a quick turnaround. This requirement is in line with efforts of the US Air Force to contain the turnaround time for quick simulation evaluations to 90 days [42].

Following all these requirements it emerged that a hybrid approach, rather than a single technique was the right tool. A large experimental simulation will be used as a starting point, but differently from conventional uses of Design of Experiment (DoE) it will employ an iterative algorithm to adapt the DoE to the evolving scenario. This way it is possible to look at the broad spectrum of the technologies of interest, while being able to adapt them to the scenarios in an iterative way.

Assumption 3 it was mentioned that the choice for this study will fall into a hybrid approach in which a DoE will be augmented by an iterative algorithm. This was chosen to capture the benefits from both approaches while minimizing issues as DoEs being extremely large and static.

Looking at the different alternatives meant to answer Following A second requirement is adaptability. There is little interest in selecting a static pool of technologies; on the contrary, it is much more interesting to be able to modify that pool so that the fleet can advance even in losing scenarios. This means that the initial set of values for each technology should not be a fixed constrained, but rather a starting point.

The third requirement is the speed of the whole process. This work is by no means trying to use high fidelity modelling – which will increase computational efforts and time. On the contrary, one of the main characteristics is the interest in using a low fidelity approach to gather insightful information to help reducing the size of the design space for high fidelity modelers, which will anyway later verify whatever technology is selected. As such, the method that is chosen should be able to analyze large design spaces with a quick turnaround. This requirement is in line with efforts of the US Air Force to contain the turnaround time for quick simulation evaluations to 90 days [42].

Following all these requirements it emerged that a hybrid approach, rather than a single technique was the right tool. A large experimental simulation will be used as a starting point, but differently from conventional uses of Design of Experiment (DoE) it will employ an iterative algorithm to adapt the DoE to the evolving scenario. This way it is possible to look at the broad spectrum of the technologies of interest, while being able to adapt them to the scenarios in an iterative way.

Assumption 3 and Gap 1, the focus now shifts in identifying which type of algorithm can cover the aforementioned gap. This decision should not be taken just in the context of finding and selecting technologies, but it should also consider the whole idea of this thesis – i.e., finding naval technology to support R&T in future Vulnerable Scenario. Although the choice and requirements of selecting the type of DoE will be discussed later, the large DoE hybrid approach leads to the next research question to determine an integrable algorithm

Research Question 6 a gap started to appear. None of the options analyzed in the literature review were able to provide a solution to cover Gap 1.

Gap 2: None of the current available options can satisfy the need to compare technologies among assets in a practical, credible and quantitative way that is adaptive and looking at large uncertainty data sets at the same time.

As such, the need for a new methodology arises. This will have to enhance state-of-the-art methodology for technology trade-off and comparison with modern statistical analysis tools. The methodology will have to address and compare technological strategies to cover the need of fleets within Vulnerable Scenario in a quantitative way for technology investment purposes.

Assumption 4: A new method can enhance state-of-the-art methodologies with modern analysis tools, meant to address technology selection and trade-off for R&D investment purposes in a quantitative way.

Following the observations and the literature review reported so far – with particular emphasis on the work done by Lempert, it is believed that this new method can be created by augmenting RDM-like methodologies with a DoE in the technology selection part, and with an adaptive system – like the signpost and trigger one – in the trade-off analysis part. If this is done, then hypothesis 3 follows:

Hypothesis 3: If a new method enhancing state-of-the-art methodologies used for technologies investment discovery is created through the iterative use of agent-based modelling, PCA and PRIM , then it will be able to adaptively find sets to technologies to increse fleets' success in naval scenarios.

4.5 Monotony of Technological Effects

Very few technologies can provide benefits without costs; a faster ship will arrive at its destination sooner, but it will be heard underwater at a much greater distance too. This means that when positively increasing one parameter for one technology it is critical to check if there are adverse effects on the scenario. Otherwise, while the fleet becomes resilient for one scenario it might become vulnerable for others that were not originally considered. Ideally, it would be desirable to have technologies which impact is positive monotone. This would ensure that throughout the iterative process no new Vulnerable Scenarios are open, this issue is therefore stated in the following Research Question.

Research Question 7: What happens if technologies effects are not positive monotone?

The issue of positive monotony of technological effects can be summarized in the observation that follows:

Observation 17: If the technology effects are not strictly positive monotone, then they can have negative effect on the scenarios, opening new vulnerable ones.

If some of the technologies have negative effects which open new Vulnerable Scenarios, the question is how can this vulnerability be captured or prevented? In addition, what happens if technologies' effects are indeed not positive monotone?

To answer these questions an experiment will have to be set up. The hypothesis in the experiment is that if new fleets are made by maximizing only one technological parameter at a time, and if all these fleets are successful in the test scenarios, then all the technologies studied have monotone behaviors. A case in which all technologies are maximized is also added to verify that if technologies interact, these interactions bring only positive effects. Regarding this, it will be used to set up the following hypothesis.

Hypothesis 4: If in a naval technology strategy, a limited number of non-positive parameters are present, then their effect can be mitigated by the combined positive impact of other parameters on the overall technology strategy.

Hypothesis 4 implies that there is an interest in identifying which parameters have a positive monotonic behavior and which do not. If a few parameters are marked with such behavior, others can mitigate their effect, or alternatively – if not critical – they could be excluded from the study. The ability of some parameters to mitigate negative effects will have to be proven of course. As such, a verification step, where the efficacy of the mitigation parameters is checked, might be necessary if negative parameters are discovered. If the all-maxed case fails, it means that when everything is maxed there are some interactions among the variables. While this might look like an issue, it is probably not, as it is not expected that any fleet will reach the level of maximizing all possible technologies.

In the end, to validate Hypothesis 4, the final fleet obtained with hybrid approach described in this chapter will be tested across all the scenarios to make sure its success rate is consistent even in the presence of some negative parameters.

CHAPTER 5. Methodology Structure

5.1 Methodology Background

Throughout the course of the previous chapters all the different pieces needed to answer the main research objective have been introduced. For each, a short literature review was provided, displaying assumptions and current gaps. In this chapter, the goal is to lay down a comprehensive methodology that, relying on the introduced pieces, will answer the research objective.

Vulnerable Scenarios and Robust Strategies are terms that have been used, and will be used, multiple times throughout this work. Even though their meaning is intuitive, it is time to formalize them. *Vulnerable Scenarios are scenarios in which the studied fleet is consistently failing a set of tasks*.

Robust Strategies, or Robust Technology Strategies, are combinations of technologies that allow the studied fleet to succeed Vulnerable Scenarios.

5.1.1 Methodology Objectives and Hypothesis

Research Objective: Develop a methodology to support trades-offs among naval assets and technologies, to assist investments on new maritime technologies in future threat scenarios

Part of the research objective is to be able to assist investments on new maritime technologies in future threat scenarios. As this implies that an understanding of those scenarios, the first part of the methodology will be devoted to that. Once those scenarios are known they can be used to frame the technological gap the fleet should cover. The second part of the methodology is dedicated to testing, in a quantitative way, how different technologies can contribute to the goal and with which proportion. Finally, once the different technology strategies are created, they are tested across the whole set of Vulnerable Scenarios to find those that are the most robust.

This methodology does not employ any high-fidelity modelling. This choice was made on purpose to make this method a tool more accessible and to reduce computational cost and complexity when scanning the whole design space. The idea behind it is that results found can be used as preliminary solutions, reducing the number of cases that high-fidelity modelers have to look at. This way, high-fidelity modelers will be able to focus on a smaller design space, making the whole investment process faster. Moreover, because of the analytical approach in the first part of the method, it is also possible to identify some Vulnerable Scenarios that might have been hidden in traditional – expert-based – studies. All of this is formalized in the following Overarching Hypothesis:

Overarching Hypothesis: If a new method is created using agent-based modelling in conjunction with an adaptive multi-objective robust design method, then, it is possible to explore hidden Vulnerable Scenarios not considered in traditional studies, and it is possible to adaptively find pools of technologies that increase the fleet's robustness in the identified Vulnerable Scenario.

5.1.2 Equivalency Conjecture

By finding Vulnerable Scenarios and Robust Strategies, the methodology is trying to emulate the same deliverable of a Capability Based Assessment (CBA). A CBA is part of the Join Capabilities Integration and Development System (JCIDS) process. CBA is an analysis in which at the end material or non-material solutions are proposed [42]. In the CBA analysts are meant to define the mission by identifying the capabilities required. This is used to detect possible gaps in the ability to solve the mission, in other words, this is used to find *Vulnerable Scenarios*. Analysts should also find solutions to cover these gaps, in other words, *Robust Strategies*. Even though CBA looks at both material and non-material solution, this thesis is limited only to the material part given the interest for investments in science and technology.

CBAs are often grounded in methods using expert elicitation. This methodology, on the other hand, tries to get the same outcome but by using low-fidelity modelling. The first advantage is that scenarios can be detached from hypotheses allowing a less biased perspective on the design space. The second advantage is that, differently from the past, technology strategies can be found in a quantitative way. This allows for a more tailored approach in finding investments that can cover the identified Vulnerable Scenarios. Unfortunately, due to the sensitive nature of the topic and to very limited access to technical information on assets, technologies and missions, formal validation is not possible. In lieu, a validation of the methodology is done by the following equivalency conjecture considering the similarities between CBA deliverable and the deliverable of this work.

Equivalency Conjecture: It is conjectured that if the Overarching Hypothesis is validated the new methodology created will be able to provide better recommendation on investments in science and technology for long term planning.

5.1.3 Methodology Structure

As mentioned in the introduction chapter, the methodology has been divided into three parts, displayed in Figure 13, to satisfy all the needs established in the research objective.

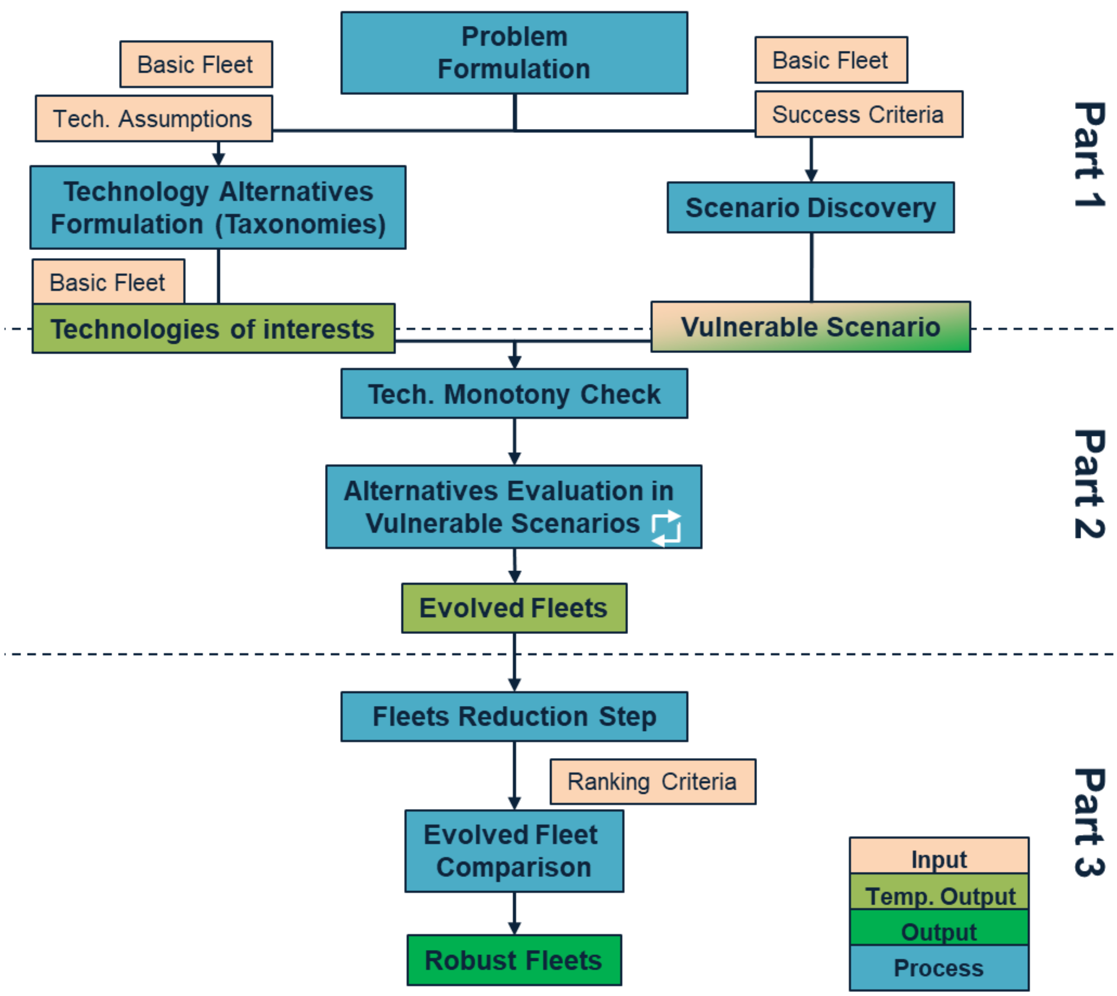


Figure 13: Proposed methodology used to achieve the research objective.

Part 1 is dedicated to finding Vulnerable Scenarios using information on the studied fleet, on the mission of interest and knowing what the success criteria of the fleet are. In this first part, only a small portion of the general DoE is used – the portion dedicated to the scenario variables. While in reality the set of Vulnerable Scenarios will be represented as a list and not in matrix form, Figure 9 helps in describing the flow of information across the different parts of the methodology.

	Vulnerable Scenario 1	Vulnerable Scenario 2	Vulnerable Scenario 3	Vulnerable Scenario ...	Vulnerable Scenario N

Figure 14: Matrix representation of Part 1 deliverable

In Part 2, information from the first part – namely the Vulnerable Scenarios – is merged with a set of technologies of interest. In this part an iterative algorithm is used to update the modelling parameters of the technologies of interest until a combination that can make the fleet succeed the scenario is found. Because of number of repetitions multiple strategies will be found per each vulnerable scenario, therefore, a way of selecting only a single strategy per scenario must be found.

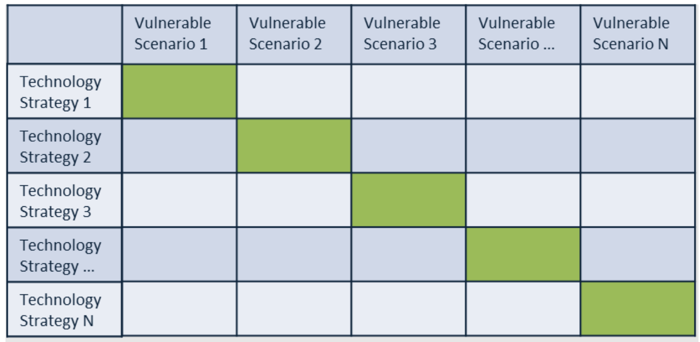


Figure 15: Matrix representation of Part 2 deliverable

Finally, in Part 3 all the robust strategies are tested in all the Vulnerable Scenario to find which one will help the fleet overcoming the greatest number of scenarios. Robust Strategies will be ranked using different criteria as costs and number of technologies invested in. If in the selected Robust Strategies, technologies with negative effects were used, a test will have to be conducted to gauge the effect of those technologies on the resulting fleets. If technologies with negative effects are present, the strategy will have to be tested also on the non-Vulnerable Scenarios to make sure that new critical cases were not created.

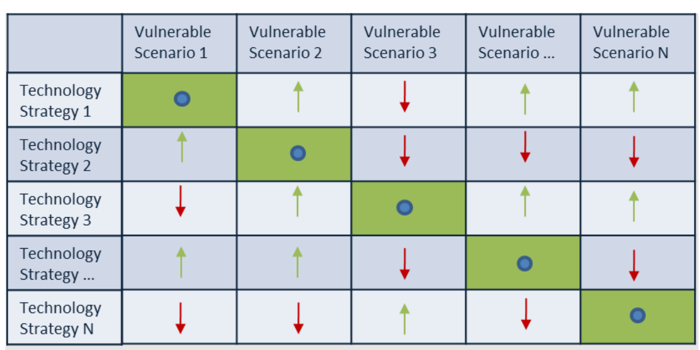


Figure 16: Matrix representation of Part 3 deliverable

5.1.4 Methodology Limitations

Before going into the details of each part of the methodology it is important to highlight the limitations that this methodology has. Following the structure showed in Figure 13, the first limitation is on the taxonomies. In fact, as mentioned before, the quality of the taxonomy used relies on subject matter experts to ensure that inputs are relevant and modelled to an extent that allows a meaningful work.

Limitations on the Scenario Discovery part are driven by the use of PRIM. For instance, since PRIM can only work with binary success criteria it cannot account for partial successes, but it requires that all success conditions are aggregated into a single one. Moreover, the size of the PRIM box affects how many Vulnerable Scenarios are captured, therefore if decision makers do not account for enough computational powers results might be limited.

Regarding Part 2, there are limitations in the iterative algorithm, the higher the number of parameters the more iterations will be required to find solutions. In this sense, while this methodology is suited for 10 to 15 technologies to be tested, if the number of technologies starts increasing too much other options (e.g., DoEs) might provide faster solutions. Furthermore, using this type of algorithms require that the taxonomy linking the technologies is complex enough so that meaningful interactions can be captured.

Limitations in Part 3 are shown in the fleets reduction step. In this step the number of fleets is reduced by aggregating them or by selecting only non-dominated ones, as it will be shown later. While this step is needed to reduce the number of fleets to a manageable one, it also implies that a lot of sub optimal solutions are discarded. This means that the fleets identified as robust at the end of the methodology will be the most expensive ones, as cheaper solutions are not considered for comparison after Part 2.

5.2 Part 1: Finding Vulnerable Scenarios

The goal of this part is to set up the whole methodology and to find the first deliverable: Vulnerable Scenarios. In this part the problem will be formulated by gathering and pre-processing important information regarding the scenarios and the assets involved. This will be mostly done in the first step called Problem Formulation. The second step will use some of the deliverables from the first one to identify which scenarios are going to be the most critical ones for the fleet of interest.

5.2.1 Problem Formulation

The first step in the methodology is problem formulation. This step follows the XLRM Framework [89] which is used to find the relevant exogenous uncertainties, policy levers, relationships and metrics of the problem.

Conjecture 2: The XLRM framework will be used to harvest the initial set of information.

Exogenous uncertainties represent all those characteristics of the problem that are unknown and that are used as variables in step two – Scenario Discovery. These unknowns are the variability on the enemy side. Examples of these variables are the enemy's aggressiveness level, weapons' efficiency, tactics, assets positions, number of assets deployed, etc. The role of these variables is to create different mission tasks to test the fleet under a wide set of scenarios.

Policy Levers are different options over which the designer has control; these are used to simulate different technologies that could be infused in a ship. Policy Levers are what we have called so far technology strategies – which are not yet robust technology strategies. Different technologies can be bundled together, with certain weights, in investment strategies that will be tested throughout the experiment. Some of the Policy Levers that can be considered are the following: new engines to reduce noise and increase speed, new sonar or radar for more detection range and more precision at range, new torpedoes or missiles for higher probability of kill, new battle management system for optimization of firing power, new Close-In Weapon System to enhance ship's survivability.

Both exogenous uncertainties and technology strategies are placed in the same DoE at the beginning of the process. The DoE will have one part dedicated to uncertainties and a second part dedicated to technologies. The exogenous uncertainty part will be a Latin Hypercube Design (LHD), while the technologies part will have all values fixed on constants that represent the current state of that technology. In Part 1 only the first part of the DoE will be used to find the Vulnerable Scenarios. In Part 2 the DoE will be screened to retain only the DoE values pertaining to the Vulnerable Scenarios which will be fixed, the technology side on the other hand will be used to find the Robust Strategies. The reason to pick use such combination of constants and variables is that there is the need to work on the same file in two different steps. Moreover, as will be explained in Part 2, the section including the constants has to be upgradable and modifiable. All of these drove the decision to use such a system. Regarding the decision to use an LHD, the need to sample inside the design space was clear. Plus, since in LHD the number of samples are independent from the number of dimensions, this enabled more variables to be tested without increasing computational costs.

Conjecture 3: A Latin Hypercube Design can be used as DoE to cover the full scenario design space of interest. These DoE can be augmented by adding constant technology columns.

The third part of the XLRM framework is relationships. This part includes all the different relationships present in the model. In general, to manage relationships there are no requirements for the software needed; in this work though – as highlighted in Chapter 3 – the choice of using an agent-based modeler was made, this is the first requirement on the tool to be chosen. Secondly, when looking at different agent-based modelers, it appeared clear that many had closed architectures that did not allow the user to create new maps or to superimpose maps to already existing bathymetric profiles. This was not acceptable in the case of this thesis as the scenario used a fictitious map, therefore a tool with an open architecture that allowed coding of new maps was required. The third requirement was established in regards of assets and technologies. Since the methodology is looking at establishing a general framework for trade-off, it is important that the tool chosen allows for easy modification of already coded assets and technologies, to enable quick re-coding if different options have to be tested, while remaining technology agnostic. While the methodology remains agnostic on the tools used, a tool called JANUS was chosen among different commercial and academic solutions. Since JANUS respected all the requirements it was chosen to be used across all the experiments in the following chapters. JANUS is an unrestricted agent-based Software Development Kit (SDK) created in house by the Aerospace System Design Laboratory and previously known as UV-core. It is Java based and it relies on the digital global environment provided by NASA WorldWinds. JANUS can be used to model complex problems and to simulate full missions, in here, different assets are defined, and different strategies and tactics are coded. The RoE and the different assumptions of the scenario are also represented and coded in JANUS. To summarize, this SDK was chosen following the need to use agent-based modelling and because of its open architecture that allowed for easier coding of new assets and technologies.

Throughout the simulation, metrics – the M in "XLRM" – will be tracked to verify if the investment strategy is robust or not and to what degree. Metrics that will be yielded are different in each use case. In general, it is expected those metrics to be parameters like mission competition time, assets lost, survived attacks, enemy destroyed etc. The way metrics are chosen is by identifying individual parameters that contribute to shaping the decision tree, these are then flagged and assigned to individual branches of the decision tree. By doing so each metric is responsible for activating or deactivating certain technologies that will be used to make Robust Strategies.

The outputs of this step are the following: the definition of a basic fleet, mission success criteria, scenario assumptions, and technology assumptions. The first 3 outputs are used again in Part 1 of the methodology, while all four of them will be used in Part 2. A term that will be largely used in the next sections is *basic fleet*, this type of fleet represents the state-of-the-art available today in terms of ships technologies, types, and numbers.

5.2.2 *Scenario Discovery*

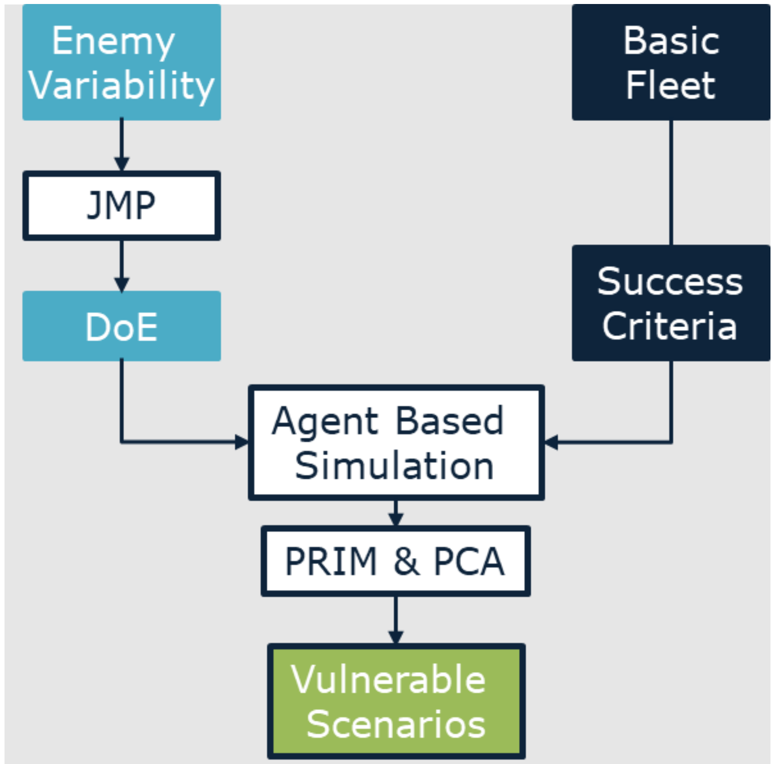


Figure 17: Scenario Discovery step diagram

The Scenario Discovery step begins with the basic information from the problem formulation one. The idea behind scenario discovery is to find which scenarios are the most critical ones for the fleet. With this objective in mind, one of the first questions that came up was regarding the level of analysis needed for the methodology to succeed. The choice fell on looking at a high-level at the whole mission.

Conjecture 4: High-level mission analysis will provide enough information to the study to make it successful

Given the difficulty in accessing detailed information due to its classified nature, the lower-level detail was not necessitated for a preliminary mission analysis in this research and will be kept at a higher level with fewer components. Moreover, a high level of details will imply a much longer computational time that would impact the ability of this study to provide results with a quick turnaround. As mentioned before, this methodology is positioned to support preliminary studies on future science and technology investments, as such, it is more important for us to be faster and quantitative rather than be extremely precise with the results. In fact, it is anyway expected that the results will be later reassessed by other entities.

5.2.2.1 Methodology for Scenario Discovery

The first step of Scenario Discovery is to gather all the scenario assumption to make the DoE. In previous chapters, it was mentioned a few times that the used methodology would be of the *agree-on-decision* type rather than the *agree-on-assumptions* one. While this remains true, and the scope of the work is still to find Vulnerable Scenarios after experiments are run, there is still the need to set a loose boundary for the problem. In this study, there are two types of assumptions: those that are used as variables (i.e., exogenous uncertainties) and those that are fixed. Given the high-level frame used for the mission analysis these assumptions will be at high-level as well (e.g., number and type of torpedoes in an enemy sub, sonar, or radar range, etc.). Ranges in variables' assumptions are ensured to be broad enough to account for unlikely options. While these ranges help in framing what a potential future could look like, there is no guarantee that a certain event will happen. On the other hand, by having such ranges broad enough it is possible to see if there is a concentration of adversary's efforts close to one of the boundaries, this implies a possible out-of-boundary behavior which will require further analysis. Some assumptions are considered fixed in this study; a clear example of this is geography. Vulnerable Scenarios are location dependent, but the complexity of considering geographical elements like the bathymetry of the sea, or the water salinity and temperature, will be excessive. Nevertheless, since the simulation is built on an agent-based modeler which can be further enhanced, if all these factors are needed in the future they could be added and treated as variables.

While dealing with large number of variables, it clearly emerged the need to have a tool for statistical analysis that could not only create some of the dataset but that could also perform the post processing and visualization of the data. When looking at the different options available, from Excel to Tableau, to JMP and Minitab, JMP was the only one to have a strong statistical analysis library which integrated all the functions needed for this thesis. JMP is in fact the workhorse of this thesis and it will be used several times throughout the methodology and in all the experiment.

Going back to the methodology, after all the variables are collected, and their ranges are uploaded in JMP, a DoE will be created. For this study a Latin Hypercube Design DoE was selected. The reasons why LHD was selected are multiple and go further on top of what hinted in previous paragraphs. First there is interest in sampling inside the design space and not just on the edges, secondly due to the lack of skewness in the values (i.e., values in variables are not weighted) there is no interest in focusing in a specific area but rather to uniformly sample the design space. Finally, by maximizing the minimum distance among points, the speed of the simulation is increased without losing quality.

Once the DoE is ready it should be loaded into the agent-based modeler. Here the basic fleet is tested in scenarios whose variability is controlled by the DoE. Because this is a stochastic process, with successful encounters controlled by random effects, fleets are tested in the same scenarios multiple times to check that success is due to the technologies on board and not to serendipity. In this step, IDs are assigned to cases so that it is possible to test the same scenarios again. This means that each combination of variables tested has one ID; but because of repetitions, each ID is tested multiple times. Therefore, all IDs are also linked to the random seed used in the simulation, and the seed is outputted for consistency checking. The combination of ID and seed is unique, and each identifies only one single case. The number of repetitions is selected using a technique called Bootstrapping.

When dealing with a stochastic simulation with stochastic responses, it is important to perform a certain number of repetitions to find the confidence intervals for the mean of the responses as well as to determine how the response is distributed. For each response, several repetitions must be determined, and the maximum number of repetitions needed for any important response must be run for each of the Design of Experiments cases. Through Bootstrapping, it is possible to estimate how the variance will change given the number of repetitions of the experiment. Bootstrapping can be summarized in the following five steps:

1. A large sample of the baseline case is created

2. The large sample is resampled with replacement to get many groups of samples with varying size (e.g., groups of multiple of 5, from 5 to 100)
3. For each group, the mean is calculated
4. By looking at the distribution of the mean the standard deviation of the mean is calculated
5. Standard deviation can be normalized by the mean to study the *coefficient of variation* c_V and its reciprocal, the *signal to noise ratio* SNR

SNR and c_V are the key to understanding how many repetitions are required for an accuracy level. While c_V is related to the trust on the mean results of the experiment, for computationally expensive simulations, it is important to reduce the number of repetitions as much as possible without sacrificing on the SNR. Reducing the SNR leads to an increase of the statistical error to unacceptable levels. More on Bootstrapping, including a practical example is presented in APPENDIX B. Bootstrapping Analysis

The results from the simulation are gathered in a single file. As mentioned in Chapter 3, if the results are too scattered PRIM can have issues identifying the Vulnerable Scenarios. As such it might be needed to preprocess the data to align the principal axis of the results with those of the hypercube of the design space. The second important preprocessing step is to aggregate in a single metric of success all the partial ones. This is needed because PRIM can only work with a single success metric. If there are multiple success criteria, the decision maker here as a choice: one can either aggregate all the metrics using an "AND" condition, or an "OR" condition. In the first case, all the success criteria will have to be met for PRIM to consider the mission successful; vice versa in the latter case, one satisfied condition is sufficient to consider the mission a success. Both preprocessing steps can be done in JMP, allowing the analyst to import the simulation results right where the DoE was, process and modify the data as needed, and output a new table for PRIM.

Conjecture 5: Because the defense field is conservative regarding standards and practices, a mission is considered a success only if all the primary objectives are satisfied.

The last step of Part 1 is to find the Vulnerable Scenarios using PRIM. There are only a few applications of PRIM, some run in Python others in C, despite the version used there should not be any difference in the result. TU Delft created a stable version of PRIM that runs on Python, this version was used to start crafting one that could work in the use case of this thesis. Once the simulation results are uploaded in the PRIM interface the algorithm will run and try to identify in the design space areas of interest. These will be given back to the user in the form of variable ranges, as can be seen in Figure 18. PRIM will also show the user the whole trade-off analysis. This is especially important because it is up to the decision maker to frame the results setting how many parameters can be controlled. Controlling parameters allows the decision maker to frame better the problem, tailoring the number of Vulnerable Scenarios found to the computational capabilities it can afford. The risk is that by leaving unconstrained too many parameters important scenarios are cut out.

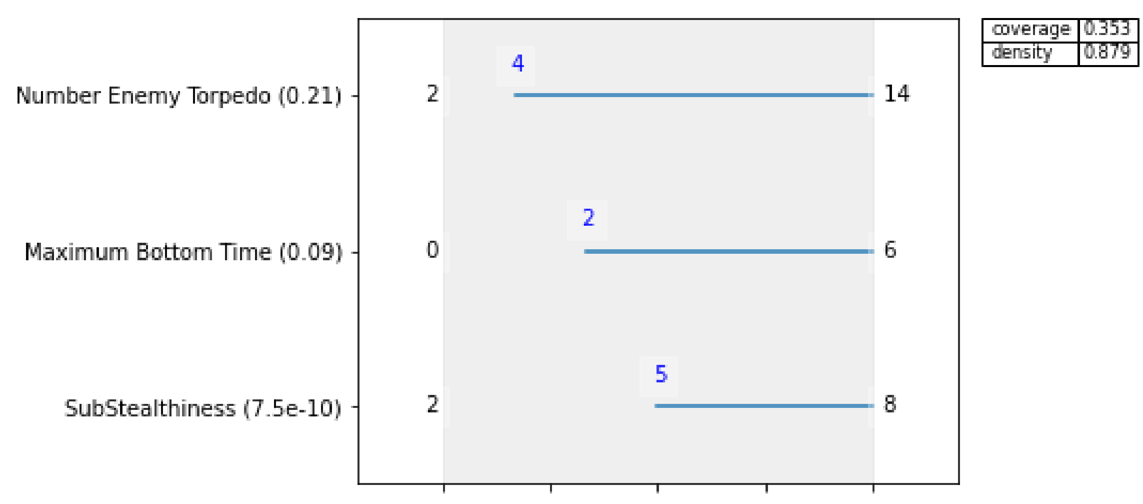


Figure 18: Examples of variables' ranges for Vulnerable Scenarios in PRIM

Once the ranges of the variables for the Vulnerable Scenarios are available these can be uploaded in JMP too. Here that information will be used to screen out only those cases that are indeed vulnerable. This step allows the user to remove not only IDs that were not critical but also cases with critical IDs whose random seed made them not critical. In the end, the Vulnerable Scenarios found are defined with two values: one ID that identifies the variables' values and one seed that is used for random parameters in the simulation.

5.3 Part 2: Finding Evolved Fleets

Once Vulnerable Scenarios have been identified it is time to look for S&T investments that will ameliorate the fleet to allow success. In Part 2 of the demonstration the focus is on technology. Starting from the information gathered in the Problem Formulation step, technologies are here decomposed using different taxonomies as described in CHAPTER 2. If necessary, a downselection of those technologies is made using approximate Pareto Frontiers. Parameters derived from the different technology of interest are then added to the DoE as constants. From here an iterative algorithm updates those values using a signpost and trigger system to adapt the fleet to the different scenarios. The output of this part is a set of evolved fleets which have proven to be successful in their scenarios.

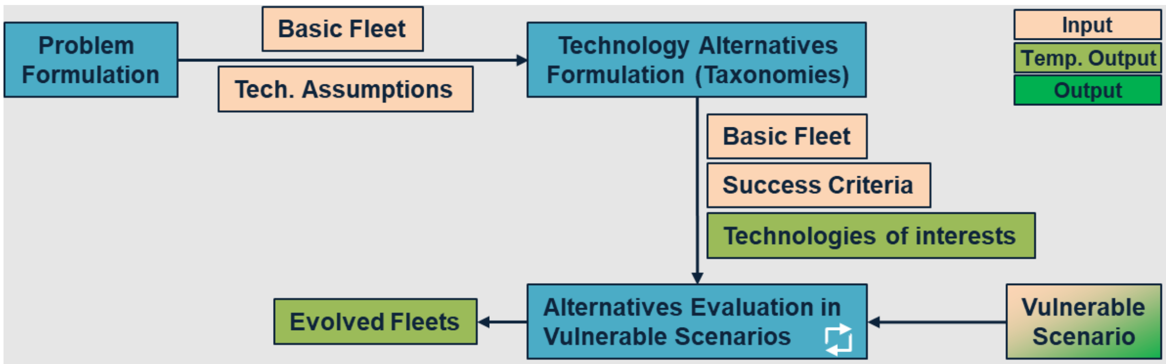


Figure 19: Schematic of the Part 2 of the methodology

5.3.1 Technology Alternatives

Technology alternatives represent opportunities for investments in the fleet. As discussed in Chapter 2, taxonomies affect how technologies are related. Knowing the taxonomy, or taxonomies, used in the model is therefore of paramount importance to understand how technologies will have to be decomposed into input parameters. All this information comes from the Problem Formulation via the technology assumption. Technology Alternatives relies on that information to define individual parameters that will affect the whole simulation and that will provide a standard for improvement.

An example of this can be the radar subsystem. Looking at the problem with a physical taxonomy in mind, the relevant parameters to be considered will be required power, noise level, required space, number of antennas, etc. These parameters will drive requirements that propagate through the whole ship design. On the other hand, using a functional taxonomy allows writing down parameters that can be readily used in a simulation. For the radar subsystem, this would be the radar range, or the radar discrimination power at range.

In this methodology, parameters derived from functional taxonomies are preferred for those components that are not creating other agents (i.e., radar, sonar, radios, etc.). For components that have a more active role and interact directly with other agents' parameters from a physical taxonomy are preferred. This distinction is made on the basis that there is an interest in parameters that can be readily used in a high-level simulation, but that are still quantitative so that can be used for S&T investments. Considering components like torpedoes, the interest is much more on the type and number as these will be the parameters driving a high-level simulation. It is worth mentioning that if the simulation focuses on something specific, the parameters of interest will also be specific, and they might be pulled from different types of taxonomies.

We mentioned before that a selection of parameters representing different technologies is added to the initial DoE, and that the value of those parameters is kept constant at what is considered the state-of-the-art value. However, depending on how many technologies the decision maker is interested in studying, a downselection process might be needed to ensure that the whole methodology remains rapid as described.

5.3.1.1 Technology of interest downselection

Taking a scenario where all uncertainties are kept constant at their most reasonable level, a Multi Objective Evolutionary Algorithm (MOEA) [88] is used to generate alternatives.

These are then tested in the chosen agent-based modeler. The goal of this step is to generate approximated Pareto frontiers to find dominated and non-dominated solutions, to then select only technologies in the non-dominated solution pools for further analysis.

Among the different MOEAs available ε- NSGAII [90][89] was selected for two reasons: t-dominance and approximate Pareto frontiers. Other MOEAs as NSGA-II [91] , Borg [92] and SAMODE [93] were discarded either because they did not provide approximate Pareto frontiers or because their dominance function was not appropriate to the problem. The t-dominance [94] utility function helps in stepping back from Pareto frontiers and focusing on approximate Pareto frontier. In fact since there is no interest in exact solutions it does not make sense to invest time and resources in finding the exact Pareto frontier. As reported in [94], t-dominance and approximate Pareto frontiers are defined as follows:

Definition 1 (t-dominance): Let $\boldsymbol{f}, \boldsymbol{g} \in \mathbb{R}^{+\boldsymbol{m}}$. Then f is said to t-dominate g for some t > 0, denoted as f $>_\varepsilon$ g, if and only if for all $\boldsymbol{i} \in \{\mathbf{1}, \dots, \boldsymbol{m}\}$

$$(\mathbf{1} + \boldsymbol{\varepsilon}) \cdot \boldsymbol{f_i} \geq \boldsymbol{g_i}$$

Definition 2 (t-approximate Pareto Set): Let $\boldsymbol{F} \subseteq \mathbb{R}^{+\boldsymbol{m}}$ be a set of vectors and t > 0. Then a set F_ε is called an t-approximate Pareto set of F, if any vector g $\in \boldsymbol{F}$ is t-dominated by at least one vector f $\in F_\varepsilon$, i.e.,

$$\forall \boldsymbol{g} \in \boldsymbol{F} : \exists \boldsymbol{f} \in \boldsymbol{F_\varepsilon} \; \boldsymbol{such\ that}\ \boldsymbol{f} >_{\boldsymbol{\varepsilon}} \boldsymbol{g}.$$

The set of all t-approximate Pareto sets of F is denoted as P_ε(F).

These two concepts are also depicted in Figure 20 and Figure 21.

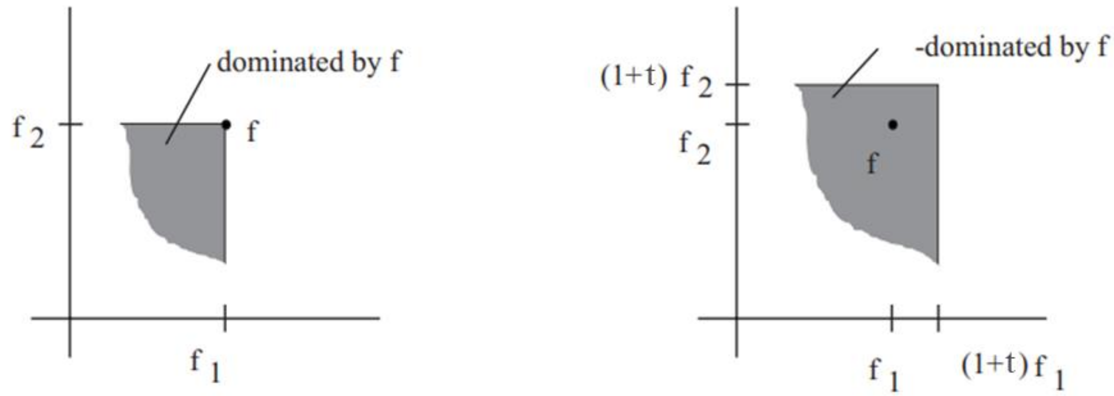


Figure 20: Difference between dominance (left) and t-dominance (right) [94]

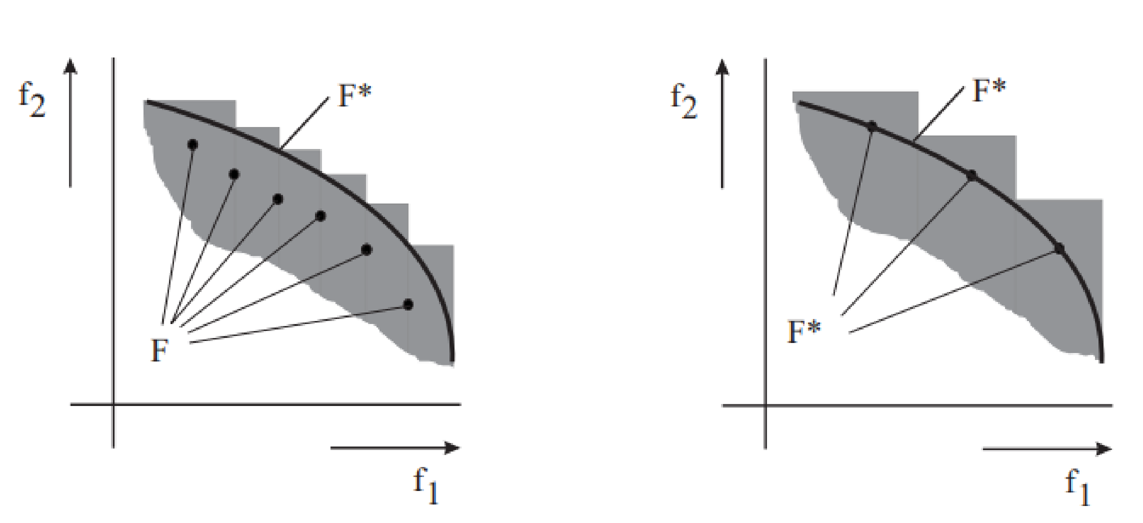


Figure 21: Difference between a t-approximated Pareto frontier (left) and a standard Pareto frontier (right) [94]

Due to the variation of the different scenarios' uncertainties considered in the first part of the methodology a full-defined Pareto frontier is not needed. Rather, this algorithm can select an area as the approximate Pareto frontier and in that area, it uses an adaptive population sizing method [95] which reduces the size of the design space while capturing all the relevant metrics. To do so, solutions are checked for density and distance. Density indicates how many solutions are similar one another – in terms of produced effect – even if they use different technologies. Distance indicates if there are niches in the design space accessible only with specific solutions.

At the end of this screening analysis, only variables present in the non-dominated solutions are put in the DoE with their value constant at the state-of-the-art level. This step is not always needed. If the number of technologies to be studied is small enough that rapid simulations are still possible, they should all be placed in the iterative algorithm.

5.3.2 Variable Monotony Check

Once the technologies have been selected, it is important to know their behavior in the simulation to avoid possible adverse effects on the fleet that will open new Vulnerable Scenarios. In fact, as discussed in Chapter 4 there is the risk of technologies negatively impacting fleets by triggering undesired consequences and effects. Among these effects there is the risk of opening of new Vulnerable Scenarios, which compromises the accuracy of what is found in Part 1 of the methodology by adding new and unknown Vulnerable Scenarios. In shot this can be summarized in the observation that follows:

Observation 18: Some technologies might negatively affect fleets by opening new Vulnerable Scenarios.

To verify if Observation 18 is correct, and if that is the case in the specific study at hand, a demonstration has been designed. This demonstration focuses on understanding the behavioral trend of each technology. Moreover, if parameters negatively affecting the fleet should be discovered, another experiment to evaluate the impact of those parameters on the overall fleet will be provided in Part 3.

From the original pool of scenarios non-Vulnerable Scenarios are taken. The hypothesis to test is the following:

Hypothesis 5: If one technology has only a positive impact on the fleet, then its use will not lead to new Vulnerable Scenarios and non-Vulnerable Scenario will remain as such.

The basic fleet can be described as $F(0) = \{0,0,0,0,0,\ldots,0\}$ where each of the variable is a technology of interest. To verify each technology, one variable at the time will be maximized and the fleet $F(0)^i$, where the i stands for the i-technology to be maximized, tested to see if Hypothesis 5 is verified or not. Since there could be possible coupling effect the fleet $F(0)^{MAX}$ in which all the technologies are maximized is also tested.

If all the non-Vulnerable Scenarios remain non-vulnerable, all is good. On the other hand, if new Vulnerable Scenarios open up they have to be addressed. First of all, the parameters that generated Vulnerable Scenario are marked. Then, if at the end of the full methodology those were used there will be another check. If they were not used no other steps are needed as all the other technologies used have only a positive impact on the fleet.

Even if some parameter, or the $F(0)^{MAX}$, produces new Vulnerable Scenario this doesn't mean that those Scenarios will be vulnerable at the end of the methodology. In fact, in this step variables are all maximized, while in the iterative algorithm they might be upgraded only to a limited extent. Consequently, this might also limit the negative effect that the parameter has on the overall fleet. Moreover, as discussed in Hypothesis 4, other parameters used in the iterative algorithm might reduce the influence of the negative ones, avoiding the creation of new Vulnerable Scenarios.

5.3.3 Alternative Evaluation in Vulnerable Scenarios

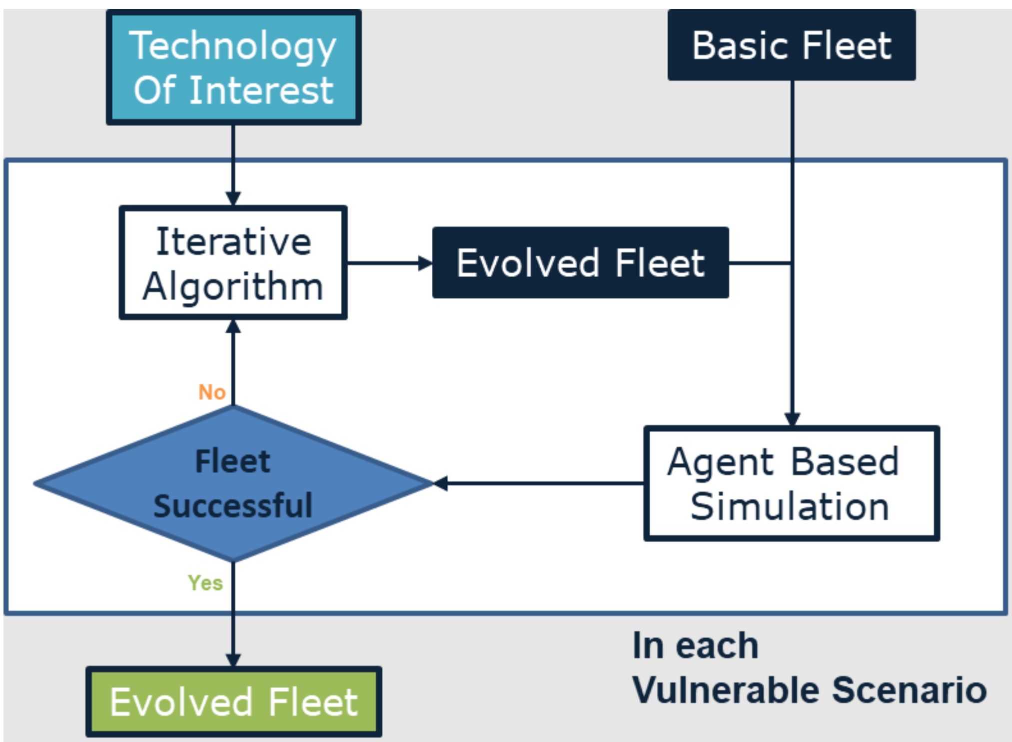


Figure 22: Graphical representation of the Alternative Evaluation in Vulnerable Scenarios process.

In this step, an iterative algorithm is used to modify the Basic Fleet by infusing technologies with different weights. This iterative step is repeated for each Vulnerable Scenario until the fleet can either succeed it, or it appears clear that the technologies chosen won't change the outcome. While ideally there would be only one evolved fleet per Vulnerable Scenario, since the simulation is stochastic for each scenario studied there are also several repetitions. This means that from each Vulnerable Scenario this step will output a series of fleets – called Evolved Fleets. These are an intermediate result as it shows fleets that overcame a Vulnerable Scenario but that have yet to prove how they perform in other scenarios. In Part 3 there will be a dedicated section to design a method to reduce the number of fleets yielded by each Vulnerable Scenario to just one.

5.3.3.1 Iterative Algorithm

The core of the Alternative Evaluation step is the iterative algorithm. The goal of this algorithm is to analyze the outputs of the simulation and rerun the same case with slightly different inputs. Inputs are changed using a signpost and trigger system. Each signpost is attached to one of the trackable metrics that are outputted after the simulation. Signposts are placed in a tree configuration where each branch is activated by the fulfillment or not of one of the main objectives. If one of the main signposts is activated, it means that somehow the fleet failed in the scenario, hence an action is triggered to upgrade the fleet before the next iteration. The four possible actions are: Reassessment, Corrective, Defensive, and Capitalizing [96].

- **Reassessment**: A Reassessment is initiated when the analysis on the scenarios shows that parameters are going out of the validity range. This causes critical assumptions to collapse and as such the investment strategy built using the policy levers might lose validity. In the Reassessment those assumptions are evaluated to see if they still hold, in that case nothing is done, or if they are violated, which stops the simulation and calls the strategy unfeasible.

- **Corrective Action**: This type of action is an adjustment to the basic strategy investment focused on preserving the outcome of the policy by using a different plan than what was originally planned. An example of this would be achieving higher speeds on an airplane by improving aerodynamics, discarding the original plan of investing in new engines.
- **Defensive Action**: This type of action is taken to reinforce the basic policy to preserve the benefits it provides. It is also used to meet outside challenges in response to specific triggers that might leave the basic investment strategy unchanged – and therefore unsuccessful. Continuing the airplane example, a defensive action would be increasing the investment in new engines to keep the original plan instead of moving to a new one.
- **Capitalizing Action**: Capitalizing actions aim at taking advantage of an opportunity to increase the outcome of the investment strategy. This action relies on the capabilities of PRIM to find regions in the design space which can be exploited. Following the same example thread, a capitalizing action will be to reduce the investment in engines because the same speed is already achieved through already exiting aerodynamic effects.

It is important to mention that these actions are implied inside the decision tree used by the algorithm, but they are not explicit. The desired effect of this step is to drive the tested strategies, iteration after iteration, toward an acceptable status in which ideally all performance metrics are above the acceptability threshold, as in Figure 23

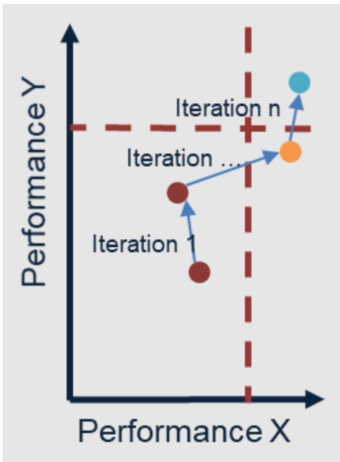


Figure 23: Desired effects of the Iterative Algorithm on tested investment strategies. In red unsuccessful strategies, in orange partially successful, in light blue successful.

The iterative algorithm is tailored to each experiment as it strictly depends on the taxonomy chosen to describe the problem at hand. However, as mentioned before several taxonomies are used in parallel to frame a complex problem properly. This means that while the variables in the algorithm are unique to the experiment at hand, if the same taxonomies used in the model are considered, and if inputs and outputs are logically connected in the model following those taxonomies, the algorithm can work independently from the specific problem.

Conjecture 6: Because the change of a technological parameter is driven univocally by specific values in the outputs, the order of technologies changed in the tree is not relevant.

The success barrier of this algorithm has two steps to ensure that the fleet is successful. The first time a fleet manages to survive and achieve all the objectives in the scenario a counter is activated. The same fleet goes again through the same scenario, if the fleet is successful again it is outputted ad Evolved Fleet, otherwise the counter is set back to zero and the fleet goes through another iteration of the algorithm.

5.4 Part 3: Finding Robust Fleets

As was defined at the beginning of this chapter, a technology strategy is considered acceptable, and therefore robust, if it is successful across a series of different scenarios as shown in Figure 24. The goal of Part 3 is to compare all the different Evolved Fleets to find which one is successful in the largest set of Vulnerable Scenarios. Each Evolved Fleet is tested in each Vulnerable Scenario, and then fleets are ranked according to their successes across scenarios. Other ranking metrics, such as budget and number of technologies upgraded are also discussed in this chapter. The first step of this part is to discuss the issue of having multiple Evolved Fleets generated by the same Vulnerable Scenario due to repetitions in the DoE.

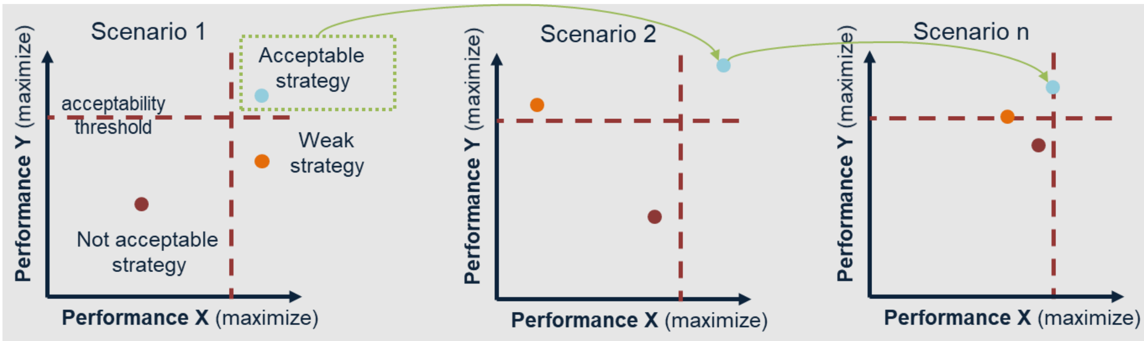


Figure 24: Representation of an acceptable Robust Strategy across several scenarios.

5.4.1 Reducing the number of fleets

One of the issues previously discussed is the high number of Evolved Fleets generated in Part 2 due to the number of repetitions present in the DoE, each of which has a different random seed and therefore is a slightly different scenario. If the number of Vulnerable Scenarios is limited the problem does not exist, and the analyst could decide to test all the found fleet in all the Vulnerable Scenarios. However, the problem escalates quickly. If there are 80 repetitions for each of the 100 Vulnerable Scenarios, assuming the same number of repetitions as needed to find the Vulnerable Scenarios, the number of simulations needed in Part 3 is 640,000. The issue of creating several fleets per Vulnerable Scenarios can be summarized as follows:

Observation 19: Because of repetitions due to the probabilistic approach, multiple Evolved Fleets are created per each tested Scenario.

Therefore, there is a need to reduce the number of fleets that will be tested in this part. In Part 2 the Basic Fleet was described as F(0) = {0,0,0,0,0,…,0} where each of the variable is a technology of interest. Generalizing that concept, the *n* Evolved fleet is described as F(n) = $\{a_1^n, a_2^n, \ \ldots, \ a_p^n\}$ where $a_i^n \in [0, M_i]$; a_i^n is the value that each technology has in the fleet *n*. There are *p* technologies in each fleet and each technology has a maximum level M_i. We assume that $F(n) > F\ (m) \leftrightarrow a_i^n \geq a_i^m \ \ \forall i$. In other words, we assume that two fleets are comparable and that one dominates the other if and only if one is superior in all technologies compared to the second fleet. In this sense it is worth asking the last Research Question.

Research Question 8: Which criteria should be used to select a reduced number of fleets to be further evaluated for robustness?

In Figure 20 the POSET structure of the Evolved Fleets generated in one single Vulnerable Scenarios is shown. In this figure, it appears that there are some dominated solutions and some non-dominated ones, the red line present connects all the non-dominated solution representing a Pareto frontier of the problem. The first screening criterion used to reduce the number of fleets is to discard all dominated solutions.

Observation 20: All dominated fleet can be discarded.

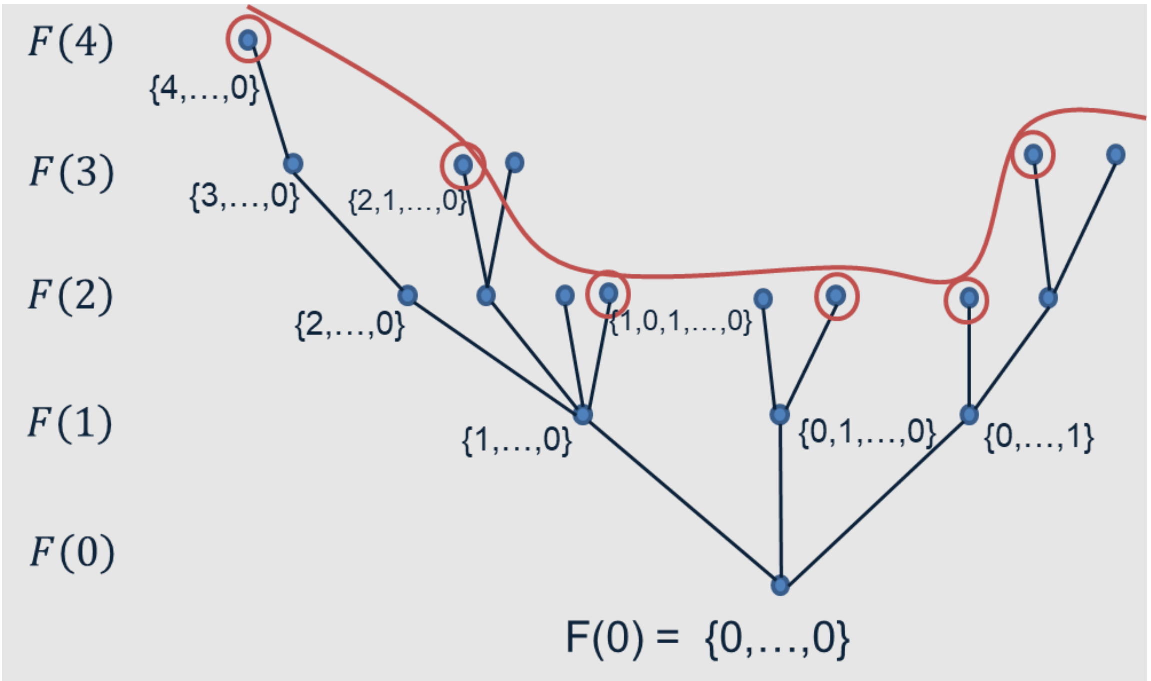


Figure 25: POSET structure of Evolved Fleets in one Vulnerable Scenario. In red circles the Evolved Fleet, the blue dots represent possible technology alternatives.

At this point there is a choice on how to further reduce the number of fleets. Still looking at Figure 25, the option is to either take all the solutions along the red line – analyzing therefore all the non-dominated strategies – or alternatively it is possible to generate one new fleet which has each technology to the maximum used by non-dominated fleets in that Vulnerable Scenario. The two alternatives are described as:

- Alternative 1 – Non-dominated fleets approach:

  In the first alternative, the idea is to test all non-dominated solutions for each scenario. The outcome is that all dominated strategies are removed but it also might lead to having multiple fleets per scenario.

- Alternative 2 – Maximized fleets approach:

In the second alternative, the number of technology strategies is reduced to one per scenario. New fleets are generated by taking the maximum of each technology in the scenario. These fleets are defined as F(t) = $\{a_1^t, a_2^t, \ \dots, \ a_p^t\}$, where $a_i^t = max[a_i^n, \dots, a_i^m]$. It is expected this alternative to be more expensive than the other as the fleets will have more technologies.

Hypothesis 6: If the maximized fleet approach is used then the fleets generated this way will dominate those generated by the non-dominated fleets approach.

5.4.2 Robust Criteria and Ranking Criteria

Before finishing the methodology by performing the comparison among Evolved Fleets the evaluation criteria to assess which fleet is robust and which is not must be defined. Following the literature in RDM, in this methodology we seek solutions which can satisfy many scenarios instead of being optimized just for one. This robustness is tested by evaluating successes of each strategy.

But how can two fleets be compared if they have the same success rate? And are there any other criteria to consider in ranking the fleets? The following approaches demonstrate some of the possible different criteria for comparison.

- Monetary approach:

  Tracking the budget is always relevant as getting a solution that works marvelously but that eats all the budget is pointless. So, a good discrimination parameter could be money, as depending on the budget available different strategies might be possible.

- Technology Variety approach:

  On the other side of the spectrum, some countries might be limited in the research they do not by money but by workers. In fact, there could be a situation in which resources are available, but the work force is lacking, setting the objective to minimize new infused technologies rather than budget.

5.4.3 Evolved Fleet Comparison

The comparison step is straightforward: each Evolved Fleet is tested in all the Vulnerable Scenarios. Successes are counted and the fleet with the most success is considered the Robust Fleet. Metrics such as budget and technology variety are also considered to check different rankings the decision maker might need.

Conjecture 7: The criteria used to rank Robust Fleets is the number of successes in different Vulnerable Scenarios. This ranking can be supported by two other discriminating criteria: budget and technology variety.

Since this is a stochastic process repetitions are needed. The same number of repetitions as in the first part of the methodology is used to maintain consistency. To ensure that one scenario does not weight more than others, successes in the same Vulnerable Scenario are summed and normalized. A zero is outputted if the success rate is not above a certain threshold, a one is outputted otherwise. This way the success rate of the fleet in each scenario becomes a binary condition; all the ones and zeros can be summed to give a numerical score representing the robustness of the fleet.

5.4.3.1 No new Vulnerable Scenarios Check

As discussed above, in some cases variables might have negative effects on the fleet. In Part 2 of this methodology, it was discussed how to identify them. If such variables were present and if they were used, one more step is needed to ensure that no new Vulnerable Scenarios are present. In general, it is not expected that any variable is brought to its maximum range, therefore even if there are negative effects it could be that these are mitigated by other variables as stated in Hypothesis 4. Table 1 describes what to do to check if the Evolved Fleets might open new Vulnerable Scenarios when the tested fleets in part 2 have negatively impacting parameters.

Table 1: Problems, Actions, and Consequences of having a variables with a negative impact on the fleet

Results from Part 2	Action	Consequence
Fleets with only one maximized technology failed	If that technology is not used – no action needed	No consequence
	If that technology is used – run fleets where that technology is present on non-Vulnerable Scenarios	Add the new Vulnerable Scenarios and adjust the ranking accordingly. No need to re check the other Evolved Fleets
Fleet with all maximized technologies failed	There are interactions among parameters – fleets should all be run on non-Vulnerable Scenarios	Add the new Vulnerable Scenarios and adjust the ranking accordingly. If this task is too computationally expensive, run only the most interesting cases
Both types of fleets failed	There are interactions among parameters and technologies with a negative impact on the fleet are present– fleets should all be run on non-Vulnerable Scenarios	Add the new Vulnerable Scenarios and adjust the ranking accordingly. If this task is too computationally expensive, run only the most interesting cases

CHAPTER 6. Finding Vulnerable Scenarios

The focus of this chapter is to be able to verify Hypothesis 2. This states that if PRIM is used together with Agent-Based modelling, then it can rapidly and credibly identify quantitative sets of naval Vulnerable Scenario. To verify the hypothesis a set of scenarios are modelled and simulated, the results are analyzed through PRIM to verify if Vulnerable Scenarios can be found or not. The outcome of this experiment will be used as a starting point for the experiments in Chapters 7, 8 and 9.

6.1 Methodology

The methodology drafted to verify Hypothesis 2 hinges on being able to simulate several scenarios using an agent-based model. Because in this experiment the focus is to find Vulnerable Scenarios the only variables studied are those related to the scenario and the environment. Fleet's variables are kept constant throughout the whole experiment.

The first step is to draft assumptions. As we saw in CHAPTER 4 Robust Decision Making techniques are considered assumptions independent as they offer the decision maker a plethora of solutions independently of the scenario assumption. This does not mean though that there are not scenario assumptions at all, but it means that there are assumption ranges rather than unique values. The first step is therefore focusing on defining the different ranges for all the variables of interest.

The second step of the methodology is looking at combining values from each variable to create scenarios. The result of this operation is a Design of Experiment. In a DoE the design space is sampled to obtain a subset of cases that will allow a broad understanding of the full model, without testing the full combinatorial space.

In step three, the DoE is transferred to the agent-based modeler where defined scenarios are run. The agent-based model provides the modelling framework for testing the fleet of interest in different scenarios.

Step four is dedicated to data filtering. In some cases, there could be a wide number of scenarios that are flagged. This would generate massive PRIM boxes leading to a too wide number of Vulnerable Scenarios. If the issue arises, this step of using PCA is applied to reduce the problem and maintain speed and continuity of processes.

The last step is the discovery of the Vulnerable Scenarios. The polished data are passed though the PRIM algorithm where a multidimensional analysis is used to identify critical regions in the design space. PRIM can find the critical region by following specific objectives as defined by the user. Objectives can include success at a given time, not losing any asset, locating objects and so on.

6.2 Experiment

To set up the experiment a scenario has been designed. In this scenario, one ship belonging to a NATO country has to perform an ASW mission to find, and if needed neutralize, 2 submarines around a group of islands. In performing this mission, different technologies and assets are tested and evaluated against the changing abilities of the enemy submarine fleet.

6.2.1 Experiment Design

To design the experiment the first step was to identify all the assets involved. For this particular experiment, the friendly fleet was made of only one frigate – based on the FREMM model – and its organic helicopter. The enemy fleet was made of two conventional submarines based on the kilo class. For each asset, relevant parameters and chosen ranges can be found in Table 2, Table 3, and Table 4.

The mission chosen for this experiment is an ASW mission. The frigate will have a certain amount of time up to 75.000 time units to locate the two submarines in a defined area – which the submarines can escape. The submarines are aggressive 50% of the times. If the submarines attack the ship or the helicopter, the RoE of the frigate is to neutralize the submarines. The mission is considered a success for the allied fleet if both submarines are located or neutralized within 30.000 time units. On the contrary, the mission fails if submarines are not located in time or if either the frigate or the helicopter is destroyed.

In this experiment the focus is to discover Vulnerable Scenarios. As such, only the enemy fleet present variables, while the NATO assets have their parameters kept constant at the most reasonable level. Some of the variables are here descripted in qualitative ways, nevertheless these have a numeric counterpart in the simulation (e.g., Hull Strength reduces the Probability of kill when the ship is hit by a torpedo by 5% per level).

Table 2: Frigate – FREMM Class Parameters

Parameter	*Value*	*Notes*
Sonar Range	35.000 m	
Sonar Quality	0	Reduces submarines stealthiness.
Torpedo Type	MU 90 lightweight	Probability of kill against submarines 0.6.
Torpedo Number	6	2 launchers.
Helicopter Number	1	
Hull Strength	Standard Plating	Probability of killing the Frigate goes from 0.55 to 0.95 depending on torpedo type.
Torpedo Decoy Quality	Standard Decoy	Decoys have a 30% probability of success.
Decoy Quantity	4	

Table 3: Helicopter – NH90 type Parameters

Parameter	*Value*	*Notes*
Dipping Sonar Range	12.000 m	
Sonar Quality	0	Reduces submarine stealthiness.
Torpedo Availability	none	
Torpedo Type	/	If available MU-90.
Sonobuoy Number	1	
Flight Endurance	10.800 time units	
Flares Number	2	Helicopter can release 2 flares charges.
Flares Quality	Standard Flares	Reduces the probablity of kill from 1 to 0.33.

Table 4: Submarine – Kilo Class Variables

Parameter	*Value Range*	*Notes*
Sonar Range	40.000 m to 45.000m	
Torpedo Type	3 qualities	Low: probability of kill vs frigate 0.55, Medium: probability of kill vs frigate 0.75, High: probability of kill vs frigate 0.95.
Torpedo Number	2 to 14	
Stealthiness	2 to 8	This represents how quiet the submarine is. The probability to be undetected goes from 20% to 80%.
Maximum Bottom Time	1.800 to 12.600 time units	This represents the time the submarine hides at the bottom of the sea trying to escape frigates.
Anti-Helicopter Missile	0 to 1	The submarine can fire against the helicopter. If available 1/3 of the torpedoes are anti-helicopter missiles.
Strategy	0 to 1	If 0 the submarine will only be passive. If 1 the submarine will attack the frigate.

6.2.1.1 Asset Behavior in the Simulation

The simulation starts with the NATO frigate entering the area of interest where the two submarines are thought to be. As soon as the frigate enters the area the helicopter is deployed. The helicopter and the frigate patrol the area using a randomized pattern, and after 1.000 time units, the helicopter deploys a sonobuoy field, which has a sonar range of 50.000 meters.

The submarines move inside a larger area than the NATO assets, this helps them hide and escape when detected. Like all the other assets, submarines move in a random pattern.

When a submarine is detected, there are different options. First, if detected by a sonobuoy and the strategy is aggressive then the submarine will destroy the sonobuoy. If a helicopter or the frigate detect the submarine shooting at torpedo, then they will attack the submarine. Second, if the frigate or a helicopter detects the submarine, and its strategy is aggressive the submarine will try to hide and re-engage the NATO asset as soon as it is not detected. Third, whenever the submarine strategy is to be passive it will try to hide as soon as detected.

When a submarine is trying to escape or hide it means that the submarine will go deeper, up to its max depth, and it will slow down to reduce noise. When the submarine reaches max depth, a timer is activated. The submarine can stay at max depth only for a certain amount of time; after the time is elapsed, it is forced to remerge. When it remerges, another timer is started, during this time if the submarine is detected it cannot escape at depth. Submarines will not attack the frigate or the helicopter when detected, therefore having a better sonar is a strong tactical advantage for the submarine.

Deployed sonobuoy fields are represented by a single sonobuoy with a variable sonar range of minimum 50.000 m. When a sonobuoy detects a submarine, it sends a message to the frigate which dispatches the closest asset to track the submarine, this could the frigate itself or the helicopter. Sonobuoys can be destroyed by submarines. If the sonobuoy is destroyed in the presence of any other NATO asset, then NATO retaliates, otherwise it is assumed that the sonobuoy just stopped working.

Helicopters fly in a random pattern in the area of interest. They can be armed with up to two light torpedoes and they can carry 1 to 5 sonobuoys. Helicopters can usually fly 10.800 time units, this can be augmented up to 21.600 time units. When a helicopter finishes its fuel it goes back to the frigate, where after a short break it is redeployed with new sonobuoys and torpedoes. The frigate can carry up to 2 helicopters.

It is important to mention that while in these paragraphs NATO assets had variable parameters (e.g., helicopter has 1 to 5 sonobuoys), this has been reported as this mission will be also used in other parts of the thesis. In this experiment NATO assets have fixed parameters (e.g., helicopter has 1 sonobuoy) as reported in found in Table 2 and Table 3.

6.2.2 Creating the DoE

All the variables from the enemy side were uploaded on JMP to create a table with many scenarios. JMP is a software for statistical analysis created by SAS Institute. This software has been widely used in many steps of this work. To create a DoE in JMP the first step was to upload the variables and their ranges in the DoE wizard. Among the different DoE options the Space Filling Design – Latin Hypercube Design option was chosen as previously anticipated. The input parameters used were those in Table 4, with whom JMP created 70 cases. These cases sampled uniformly the design space, without the need to perform a full factorial DoE which would have had 32.928 cases.

To each scenario a unique ID was assigned for tracking purposes. After conducting a bootstrapping analysis of the scenarios, the number of repetitions was set at 100. The final output of the experimentation phase is a csv file with the scenarios that can be uploaded into the agent-based modelling framework. For this case study, this has 7.000 total scenarios including repetitions.

6.2.3 JANUS

JANUS is an agent-based modelling software developed by the Aerospace System Design Laboratory (ASDL) at Georgia Tech in the past decade. JANUS is written on JAVA, and throughout the course of the author's academic career at ASDL it has been the backbone of his research.

The csv file can be uploaded in the JANUS interface. Once uploaded, the software asks if any optimization is required or if the file should be simply run in batch mode. In this experiment the scenario file was run in batch mode, without any further optimization. This meant that the software went case by case reproducing in the simulation the information contained in the CSV file. To speed up the process, visualization features were turned off. To grant repeatability, the 16-digit random seed was included in the output file for each case.

The output parameters can be found in Table 5 with a synthetic explanation of why each parameter was yielded.

Table 5: JANUS First Experiment Output Parameters

Parameter	*Note*
total Runtime	This is the total running time of each case. It can be lower than the maximum (75.000) if the simulation is stopped due to success or failure.
seed	The random seed used in the case, stored for repeatability.
Number of Frigates Killed	Number of frigates destroyed by submarines, if bigger than 0 the case is a failure.
Number of Torpedoes Fired	Number of torpedoes fired by the frigate.
Total ASW Time	Total mission time. If 0, the mission was not successfully completed.
Number of Helicopters Deployed	Number of helicopters deployed. It is equal to the number of helicopters available, unless some are destroyed.
Number of Helicopter Sorties	Number of times each helicopter has been deployed times the number of helicopters available.
Number of Sonobuoys Launched	Number of sonobuoy fields launched by helicopters.
Bad Configuration	Used to mark optimized configurations that still fail the mission. Not used in this experiment.
Mission Type	Verification parameter for optimization algorithm. Not used in this experiment.
Number of Submarines Identified	Number of submarines identified either by the frigate or by the helicopters. If the value is 2 for more the 30.000 time units, the simulation is stopped and considered a success.
Number of Submarines Killed	Number of submarines killed. If the value is 2 the simulation is stopped and considered a success.
Number of Submarines Escaped	Number of submarines escaped from any of the NATO assets.

Torpedo Jammed	Number of torpedoes jammed by the frigate decoys.
Torpedo Fired	Number of torpedoes fired by enemy submarines.
Number of Helicopters Killed	Number of helicopters killed by enemy submarines. If bigger than zero, the mission is considered a failure.
Number of Decoy Fired	Number of decoys fired by the frigate.
Number of Sonobuoy Killed	Number of sonobuoys destroyed by enemy submarines.
Scenario ID	ID of the scenario, stored for traceability.

While JANUS outputs a wide array of variables in this experiment the focus is only on 3 of them: N*umber of Frigates Killed, Total ASW Time*, *Number of Helicopters Killed*. These variables identify which scenario is a failure and which is a success. Other variables are used to draft conclusions on what are the causes of failure or success. Finally, in the outputs there are some variables that are only stored for bookkeeping purposes and for traceability.

6.3 Results

PRIM was run using the output from the JANUS simulation. In more than 80% of the cases the fleet failed the mission, proving that it was not able to satisfy the objectives. 5732 out of the 7000 simulations were therefore deemed of interest by the code which means that they were marked for furtherer analysis. The trade-off analysis showed almost a flat behavior in balancing density and coverage, as it is possible to see in Figure 26. The flat behavior is not bad, but it shows that the only trade to be made in this case is between the number of restricted dimensions and coverage. To limit the number of restricted dimensions to the highest number possible (i.e., 4), coverage was therefore sacrificed. The full result of the first PRIM iteration can be found in Table 6. It is important to mention that while cases of interest were only found in the 11% of the design space, they represent and could be used to reproduce almost 90% of the whole cases of interest.

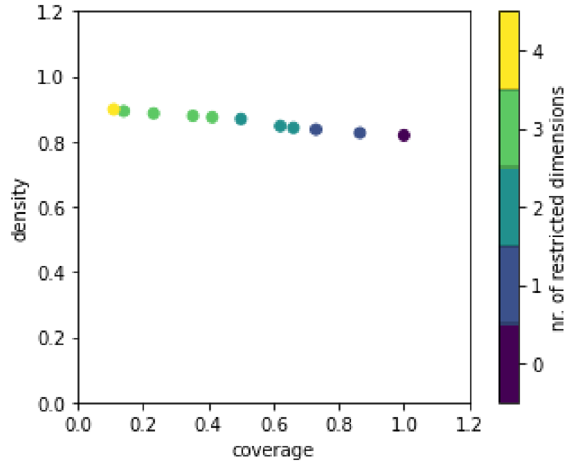


Figure 26: PRIM Trade-off analysis result

Table 6: PRIM iteration results

Mean	0.898571
Mass	0.1
Coverage	0.109734
Density	0.898571
Restricted Dimensions	4

Looking now inside the trade-off graph it is possible to expand the result for each of the cases in the picture. As was mentioned before, this allows the decision maker to accommodate the results to its ability to process, only one point is needed in the analysis. In this case, point 11 was chosen - the one on the far left in Figure 26 - as it was the one with the highest number of restricted dimensions, allowing for a more insightful comparison with data from JMP. Results of point 11 are also visible in Figure 22, where the PRIM boxes are shown in red highlighting Vulnerable Scenarios variables' intervals.

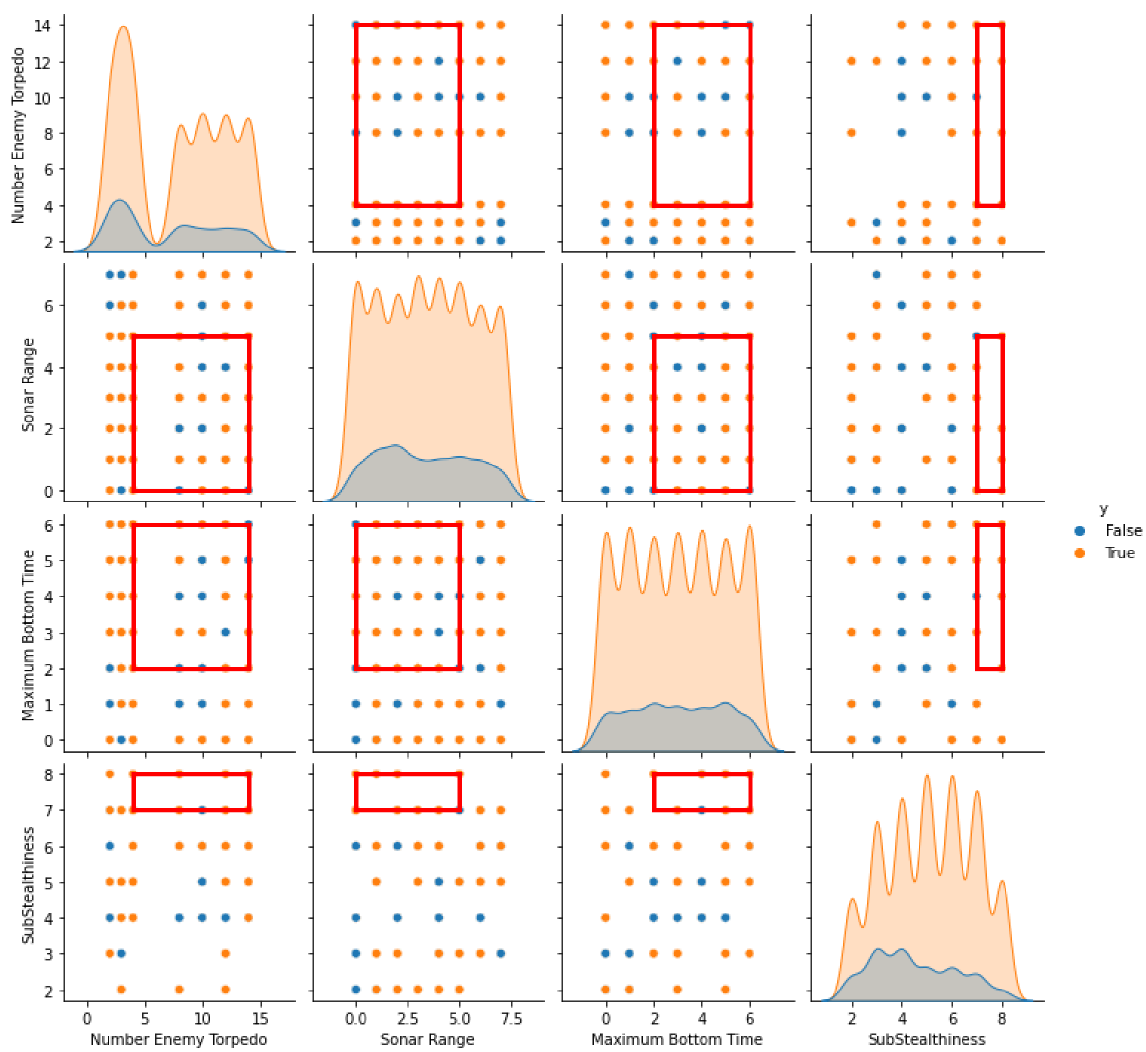


Figure 27: Results from PRIM, the red boxes represent the set of Vulnerable Scenario. In blue cases where the NATO fleet had success, in orange where it failed

To augment the result obtained by the PRIM algorithm it is possible to visualize them on JMP. Figure 28 shows the comprehensive results of the experiment, in here the maximum value per Scenario ID is 100, representing the number of repetitions. This means that a case reporting 98 had a 98% failure rate in the simulation. It appears clear from the graph that most of the scenarios failed – either because of the time needed by NATO to find the submarines or because of the enemy's ability to outperform NATO in combat. This very high failure rate highlights the need for technology investments to ensure that the fleet can successfully perform the mission assigned.

Figure 29 shows a subset of the information present in Figure 28. Here, those data were filtered using the variables ranges of the PRIM box that PRIM identified. It is important to highlight how some of the cases with a high failure rate (e.g., ID 4, 22 or 61) were not included in the subset discovered by PRIM. This is due to some of the repeated values present in those ID, which were filtered out by PRIM when it was balancing density and coverage.

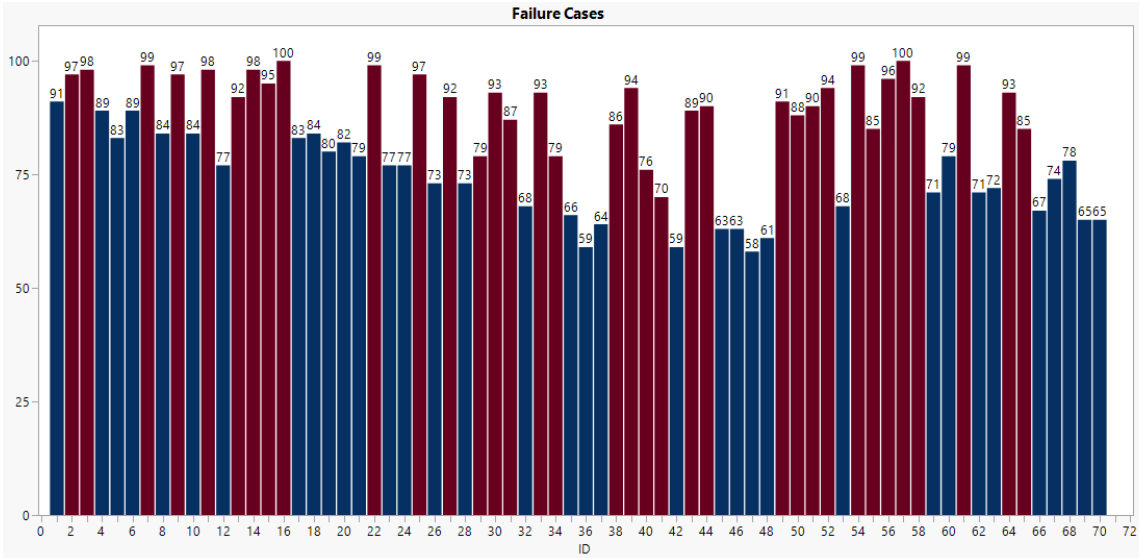


Figure 28: Failure rate by cases, red bars are aggressive deterrence strategies, blue bars are passive deterrence strategies

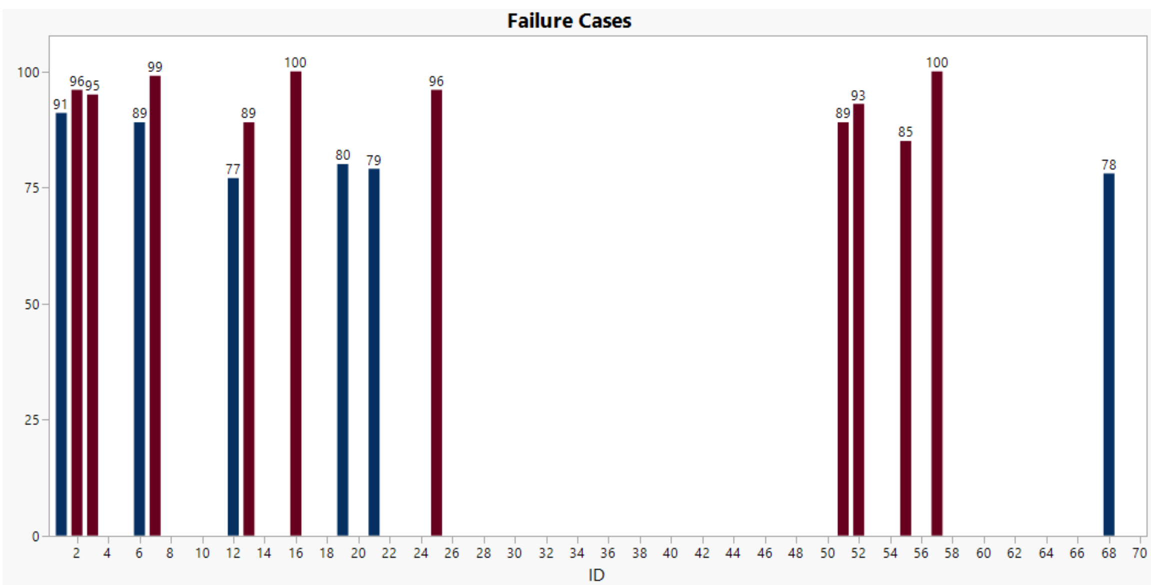


Figure 29: Subset of Vulnerable Scenarios ID identified by PRIM (4 restricted dimensions)

6.4 Conclusion

What this experiment has demonstrated is the successful proof of Hypothesis 2. To verify this hypothesis a set of naval scenarios were created on JMP and tested in JANUS. The outcomes from JANUS were later processed on JMP and using PRIM.

In performing this last comparison, it emerged that the PRIM box captured most of the Vulnerable Scenarios. This can be seen as for both deterrence strategies, the case IDs with the highest failure rate were highlighted by PRIM, as it appears in Figure 29. By using PRIM, the number of scenarios has been reduced 5 times, without losing valuable information. This was confirmed by the information on density and coverage provided in the Result section.

6.4.1 Picking a different point

As shown in Figure 26 PRIM provided a set of points that are a trade-off between coverage of the design space and density of the solution of interest. To provide the results and conclusions described in the previous section, the point studied was the one with the highest number of restricted dimensions as done by Lampert in many of his works.

While not as common, it is also possible to pick a different point. In this experiment, the trade-off showed a quick degradation of the coverage parameter with an increase of restricted dimensions. As such it is interesting to try to increase the coverage factor by picking a point that has only 3 restricted dimensions instead of 4. We will see that by doing so more Vulnerable Scenarios can be included in the study, expanding the number of scenarios above the minimum required one. This will not impact the results, but on the other hand will increase their accuracy.

Looking at Figure 26, choosing the point with the highest coverage would be point 7 in light green. Comparing this point with the previous one (i.e., point 11) it appears clear how the coverage is much higher while the density is only minimally downgraded. The range of values for common variables is much wider in point 7 seen in Figure 30 and Figure 31. This shows that since the algorithm is restricting less variables it is widening the range to capture more cases. This is confirmed by comparing the amount of cases ID captured in point 7 visible in Figure 32 with those captured in Figure 24

Results in Figure 32 are the same as those in Figure 29 plus, some more due to the higher coverage value of this point. These results show how by restricting less variables – hence by reducing the dimensionality of the PRIM box – the number of Vulnerable Scenarios increases.

Finally, it is important to underline how also this point verifies Hypothesis 2. All the scenarios found are in fact among the most critical ones. If we look at the top 10 cases for failure rate in Figure 28, 8 of those are present in Figure 32, proving once again the scenario discovery ability of this method.

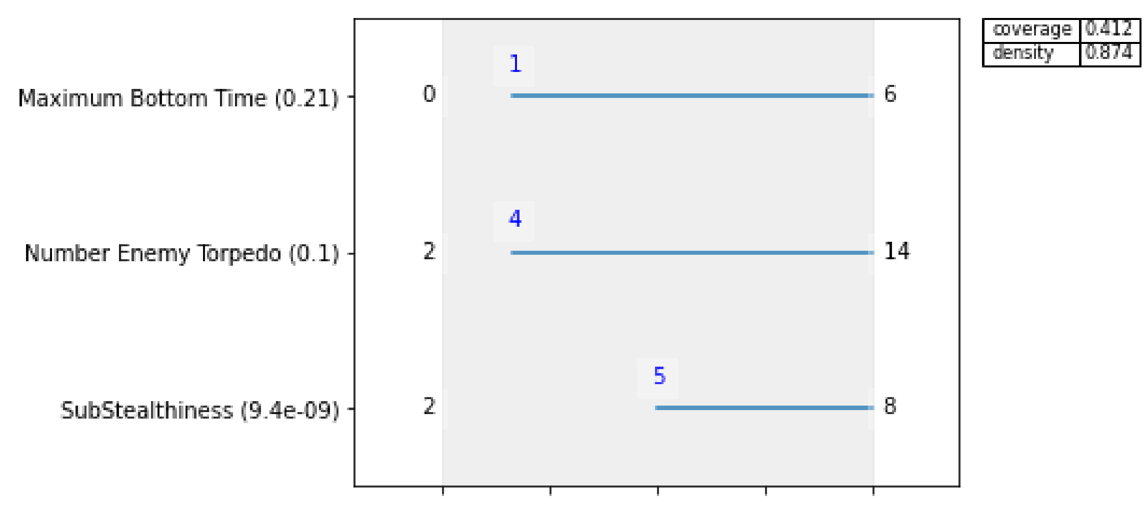


Figure 30: Ranges of each of the 3 restricted dimensions.

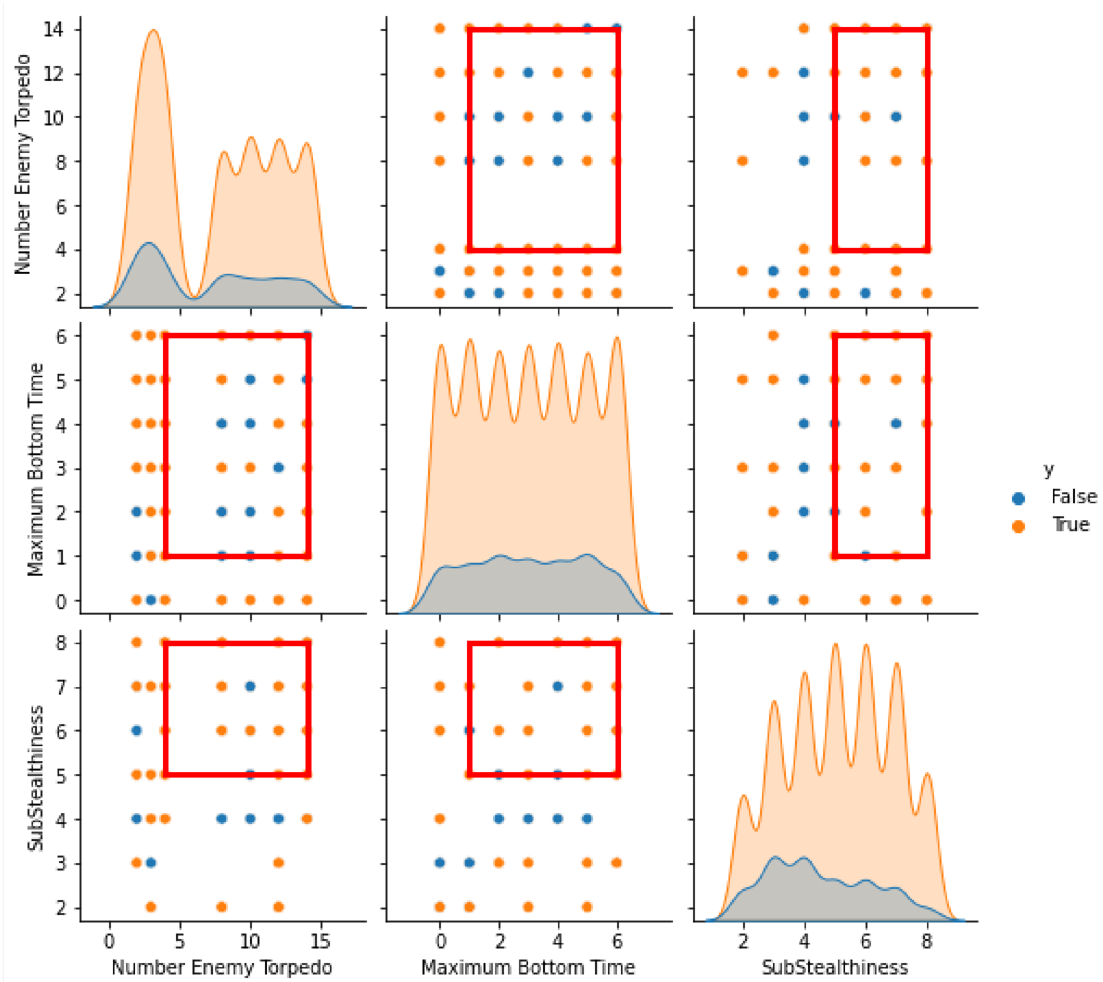


Figure 31: Results from PRIM (point 7), the red boxes represent the set of Vulnerable Scenario. In blue cases where the NATO fleet had success, in orange where it failed.

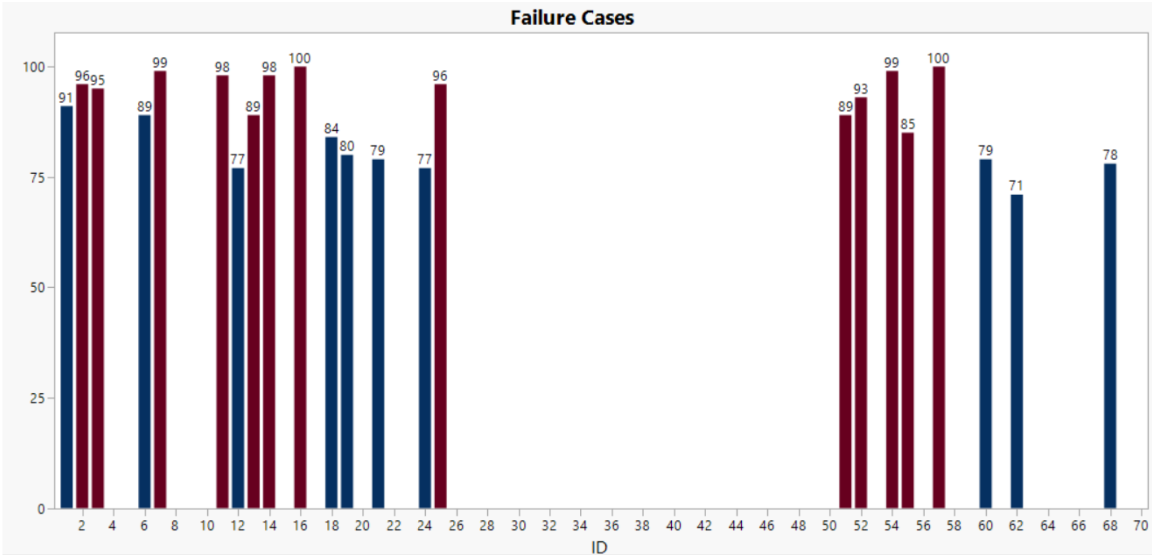


Figure 32: Subset of Vulnerable Scenarios ID identified by PRIM in point 7 (3 restricted dimensions).

CHAPTER 7. Identifying Robust Technology Strategies

As was explained in Chapter 4 current methodologies lack the ability to adaptively find sets of technologies that can quickly overcome Vulnerable Scenario. The experiment that this chapter will demonstrate aims to verify Hypothesis 3 by testing an in-house developed iterative algorithm that relies on Adaptive Robust Decision Making but goes forward by mixing it with DoE. The use of this new algorithm will also provide a complete response to Following A second requirement is adaptability. There is little interest in selecting a static pool of technologies; on the contrary, it is much more interesting to be able to modify that pool so that the fleet can advance even in losing scenarios. This means that the initial set of values for each technology should not be a fixed constrained, but rather a starting point.

The third requirement is the speed of the whole process. This work is by no means trying to use high fidelity modelling – which will increase computational efforts and time. On the contrary, one of the main characteristics is the interest in using a low fidelity approach to gather insightful information to help reducing the size of the design space for high fidelity modelers, which will anyway later verify whatever technology is selected. As such, the method that is chosen should be able to analyze large design spaces with a quick turnaround. This requirement is in line with efforts of the US Air Force to contain the turnaround time for quick simulation evaluations to 90 days [42].

Following all these requirements it emerged that a hybrid approach, rather than a single technique was the right tool. A large experimental simulation will be used as a starting point, but differently from conventional uses of Design of Experiment (DoE) it will employ an iterative algorithm to adapt the DoE to the evolving scenario. This way it is possible to look at the broad spectrum of the technologies of interest, while being able to adapt them to the scenarios in an iterative way.

Assumption 3 and Gap 1, the focus now shifts in identifying which type of algorithm can cover the aforementioned gap. This decision should not be taken just in the context of finding and selecting technologies, but it should also consider the whole idea of this thesis – i.e., finding naval technology to support R&T in future Vulnerable Scenario. Although the choice and requirements of selecting the type of DoE will be discussed later, the large DoE hybrid approach leads to the next research question to determine an integrable algorithm

Research Question 6.

Hypothesis 3 states that if a new method that will enhance state-of-the-art methodologies with modern analysis tools is created, then it will be able to adaptively find sets to technologies to increse fleets' success in naval scenarios. This means that if the hypothesis is verified, from the original Basic Fleet it will be possible to create a series of new fleets using modern analysis tools (i.e., DoE in cunjunction with the iterative algorithm).

While this experiment can be fully disconnected from the one conducted in Chapter 6, it makes sense to rely on that testbed to maintain continuity of the case study. The information taken from that experiment include: the NATO fleet composition, the asset’s behavior and assumptions, and the discovered Vulnerable Scenarios identified via PRIM. It is important to highlight that this experiment could be conducted with any scenario, vulnerable or not, and with any set of assets.

The robust technology strategy experiment was created to test and evaluate the relative success rate increase from the Basic Fleet to the Evolved Fleets. By quantitatively evaluating success rate it is possible to verify if Hypothesis 3 stands or not. In this sense, this experiment was chosen because it enables to test the Adaptive Robust Decision Making Algorithm created, and the method used to find Evolved Fleets.

7.1 Methodology

The methodology follows that of Figure 22, demonstrated on the case study of the previously selected scenario from the vulnerability experiment in Chapter 6. Once initial conditions, success criteria and the number of repetitions are chosen the assets can be tested. All the information regarding the scenario and its variability is present in the DoE, which, if the analyst was not provided already, will have to make. In making the DoE the analyst should capture all the relevant parameters that are present in the vulnerable scenarios to be studied. If already made Vulnerable Scenarios are used, then it is critical to make sure that the random seed that governs stochastic events is present. Without this information whatever is discovered throughout this methodology will not be repeatable. Moreover, in this case repetitions are not needed as the seed will keep the scenario constant, otherwise bootstrapping can always be used to find the number of repetitions needed. For result consistency all the cases related to the ID of the Vulnerable Scenario should be run. Regarding the technology of interests, these should be present in both the DoE and in the iterative algorithm. The technology values in the DoE are used as a starting point for the algorithm.

Once the full set of assumptions is ready it should be loaded in the chosen modeler. To ensure consistency, in this work it was chosen to use the same agent-based modeler used for the previous experiment (JANUS). At this point the Basic Fleet is tested in each scenario. When the fleet is not successful it goes through an iterative algorithm that increments or decrements a subset of the technologies of interests. This subset is chosen depending on the previous simulation output.

The iterative algorithm employs a signpost and trigger system that enables us to adaptively select different technologies to be used in the next simulation. This system is designed to iterate the fleet until satisfaction of the mission minimum requirements. If there are any optional objectives the algorithm will try to satisfy those as well, but once the main objectives are achieved the fleet will be output and the iterative process stops. This algorithm must be tailored to the use case at hand, but the logic it follows does not change through use cases.

To activate the algorithm one of the main objectives of the mission must be to fail. The failure of some mission criteria is used as a signpost to activate different branches of the algorithm. For every objective that is not satisfied the algorithm should have a branch of technologies that will help cover that gap. Different signposts activate different branches, and multiple branches can be activated during the same iteration. This way fewer iterations are needed to find the evolved fleet as the technology search in the algorithm is parallelized. To select which technologies to update inside the branches the algorithm checks the output of the simulation looking for technologies that can cover that gap. If there are multiple technologies, they should be arranged using a criterion – in the case of this work the criteria chosen was the estimated total cost. This means that technologies were arranged from the cheapest to the most expensive one keeping in mind not just the estimated cost of integration, but also of research and development, and operation. This criterion was chosen as this work is dedicated to satisfying a need for Science and Technology investment. When all the success criteria are satisfied the fleet can be output and the algorithm stops. A maximum number of iterations should be set up to avoid cases in which the fleet keeps iterating without being successful. A good measure for identifying this cap is to calculate the maximum number of iterations needed to reach the bottom of a branch and added a margin of error to account for branches alternation.

These outputs should then be loaded on JMP to perform an analysis of which technologies provided the most success, which scenarios were unrecoverable with the current set of technologies, and the number of iterations needed. This last piece of information is the key one to provide to decision makers for them to make an accurate evaluation of investments in the technologies of interest.

7.1.1 Methodology and Experiment Validation Against Hypothesis

This methodology is the one described in Assumption 4. The reason this experiment was chosen is that on one hand it is S&T oriented, and on the other it allows extending already existing methodologies to the naval field through modern analysis tools. It provides a good framework to test an iterative approach that allows the user to reduce the number of iterations needed to find how different technologies perform, reducing consequently the entire process time. Looking at Hypothesis 3, this experiment provides enough information to validate, or discard, the hypothesis. If the experiment is successful then it will be visible by comparing results against what was found in Chapter 6, more specifically in Figure 32. Moreover, to assess the speed of the whole methodology it is possible to compare the number of total runs – number of cases times the median number of iterations – against the total number of cases using just a DoE – times the number of repetitions needed.

As described in Chapter 4 it might be possible that some technologies of interest are not positive monotone, but they negatively affect the fleet opening of new Vulnerable Scenarios, and de facto compromising this iterative step. To address the compromised state, an experimentation that will be addressed in Chapter 8 closes the gap created by these results.

7.2 Experiment

The experiment follows the one already performed in Chapter 6. The allied fleet is composed of one frigate – based on the FREMM model – and its organic assets. The enemy fleet was made of two conventional submarines based on the kilo class. For each asset, relevant parameters and chosen ranges can be found in found in Table 2, Table 3, and Table 4. Table 2 and Table 3 provide the initial values used by the iterative algorithm, ranges that will be tested are presented in Table 7and Table 8. The enemy fleet is still based on Table 4 but the Vulnerable Scenarios chosen are those visible in Figure 32.

Table 7: Frigate's Variables Range

Frigate – FREMM Class	*Min*	*Max*
Sonar Range	35.000 m	42.000m
Sonar Quality	Increase probability of detection of enemy subs by 0.	Increase probability of detection of enemy subs by 0.8.
Torpedo Type	Light Torpedo with probability of kill 0.6	Heavy Torpedo probability of kill 0.8
Torpedo Number	6	12
Helicopter Number	1	2
Hull Strength	Probability of kill against frigate reduced by 0.	Probability of kill against frigate reduced by 0.5.
Torpedo Decoy Quality	Success rate 0.3	Success rate 0.8
Decoys Quantity	4	10

Table 8: Helicopter's Variables Range

Helicopter – NH90 type	*Min*	*Max*
Dipping Sonar Range	12.000 m	15.000 m
Sonar Quality	Increase probability of detection of enemy subs by 0	Increase probability of detection of enemy subs by 0.8
Torpedo Availability	Not Available	2 torpedoes available
Torpedo Type	Light Torpedo probability of kill 0.6	Heavy Torpedo probability of kill 0.8
Sonobuoy Fields Number	1	3
Flight Endurance	10.800 time units	32.400 time units
Flares Number	2	4
Flares Quality	Probability of kill against helicopter reduced to 0.33	Probability of kill against helicopter reduced to 0.17

7.2.1 Experiment Process

The experiment starts by uploading the DoE in JANUS. The DoE includes all the Vulnerable Scenarios identified in Chapter 6, whose IDs are shown in Figure 32. For all these IDs, all the 100 repetitions were included, and the random seeds generated in the previous scenario was added to the DoE to ensure that JANUS simulated the same scenario. The total number of cases was 2300.

At the end of each run JANUS checks if the mission is successful, which means that all 3 success conditions are met: no frigate is lost, no helicopters are lost, and submarines are either located or neutralized before 30.000 time units. If one of these conditions is not met, then JANUS activates the iterative algorithm part. During the iterative algorithm part, the code checks some outputs that are tracked (signpost) and depending on their values it triggers an action, more on this in the next section. The iteration cap was set to 50, which means that as the algorithm tries to upgrade the fleet 50 times it will quits at the maximum label it as “bad configuration”. This number was chosen by calculating the maximum number of iterations possible combining all the technologies of interest and adding to it a 20% to account for situations in which branches are not optimized at the same time.

Once all the fleets have been optimized results are outputted and reuploaded on JMP to evaluate them. The interest here is in understanding which technology worked and which did not.

For all the fleets that were optimized and were successful in their relative scenario we study the combinations of technologies to assess which ones were mostly used and which ones were less used. The level of use is also important: did the algorithm upgrade one technology to its maximum, or several to a lower level? What is the number of technologies that were infused in the end? What is the cost of the evolved fleet? All these questions must be answered in the results to be able to accurately assess the results obtained.

For those fleets that failed after 50 iterations it is important to understand why. It could be that the set of proposed technologies was not tailored to the task, or it could be that the ranges set up for the variables were not high enough to allow a mission accomplished. It is also important to study the scenarios in which these fleets are failing. While these are still Vulnerable Scenario, they are also scenarios for which we are not preparing yet (or at least we have not tested the right technologies for them).

7.2.2 Iterative Algorithm Structure

For this experiment the iterative algorithm was set up to have 3 branches, one per each of the success conditions. The first branch – the one regarding the survival of the frigate – is reported in Figure 33. The other two, survival of the helicopter and maximum ASW time, are reported in Appendix D. The outputs that are used to trigger actions are of two distinct categories: those regarding the NATO asset status at the end of the task and those regarding the enemy status at the end of the task.

Regarding the first type of outputs, these are information like how many torpedoes are left, or what type of torpedoes was used. For the second type of output the information are focused on what could NATO know after the engagement, i.e., if the submarine is still alive, if it escaped and if it was aggressive or not. The logic is to evaluate what happened using the second type of information and then to look at the first type to make it better next time.

Increments in the variables are done by pushing the technology to the median value first, and then after all other technologies in the same branch have been tested, the remaining part of the design space is unlocked. This of course was a design choice which could be changed depending on the need of the study. It is conjectured that if there are enough workers to engage in different technologies the budget-optimum logic will follow this approach. On the other hand, if the decision maker has to pick only one technology to finance, it could be interesting to see how an approach in which technologies are maxed independently will work.

Conjecture 8: It is conjectured that budget-optimum logic will push one technology to half of its potential and move to the next one, than pushing one technology to the extreme before testing something different.

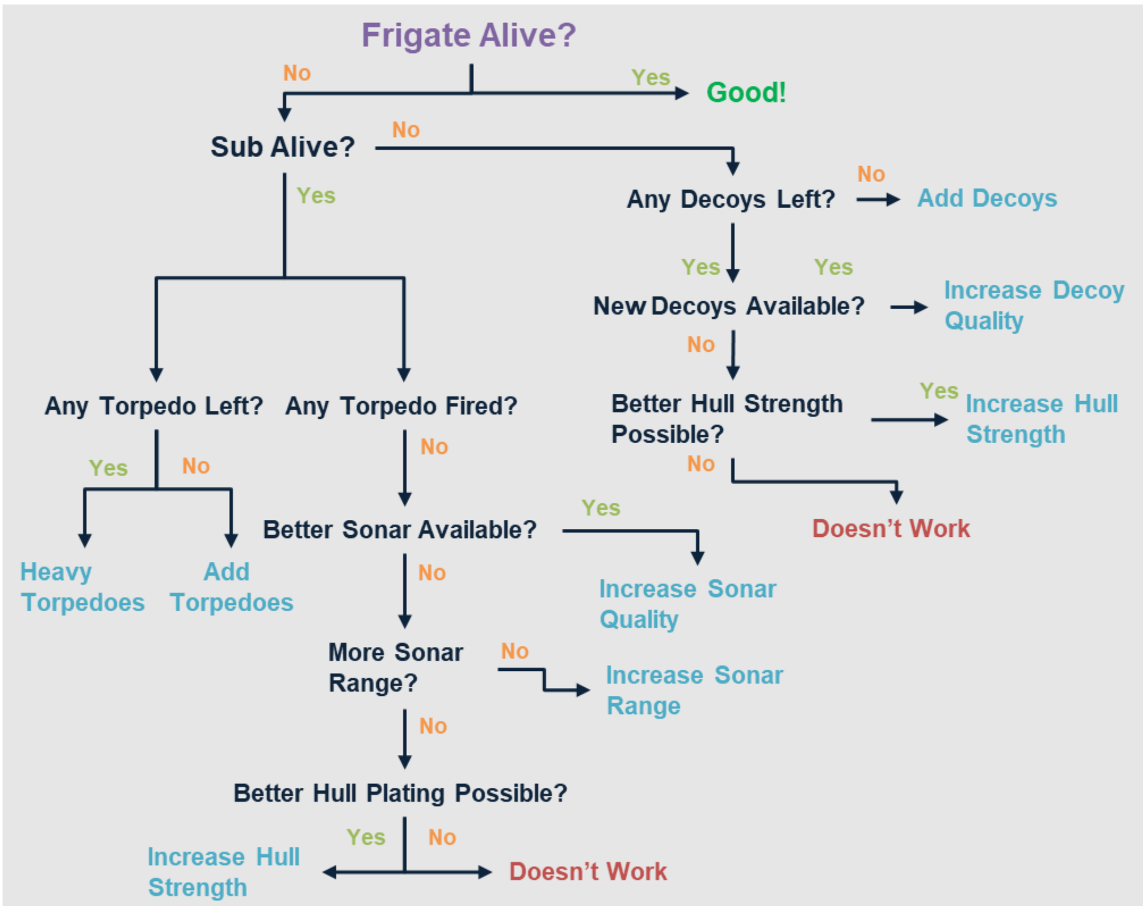


Figure 33: Iterative algorithm structure for one success condition in experiment 2

In the algorithms as the one depicted in Figure 33 it is possible to end up not solving the problem. In fact, even if many technologies are tested it is possible that the pool of technologies of interest is not up to the task. In those cases, when the algorithm reaches one of the "doesn't work" statements, configurations are marked as "Bad Configurations" so that analysts know that what was provided does not meet the minimum requirements of the mission.

7.3 Results

Figure 34 shows how many cases are still vulnerable. By looking at it, it appears how almost all the Vulnerable Scenario have been succeeded by the Evolved Fleets. There are only few exceptions as IDs 57, 25 and 16 in which the failure rate is still high 49, 39 and 30% respectively. Expectedly, the hardest cases to resolve were those in which an aggressive strategy was employed by enemy forces. This is manifested by the lack of any passive deterrence strategy cases in Figure 34.

The fact that there are still unresolved cases should not surprise the reader as the success of this algorithm is dependent on the technologies that are tested in the Vulnerable Scenarios of interest. In this sense, Figure 34 tells us that different technologies are still needed as the ones provided are not sufficient to solve all cases.

It is important to underline the fact that these technology investments are the key for fleet's success in Vulnerable Scenarios. This emerges clearly if the results of this experiment, as shown in Figure 34, are compared with those found in the previous chapter and shown in Figure 29.

The other relevant piece of information in this paragraph is the level at which each technology was used. The contour plots of the technologies used are in Figure 35. In this diagram each row represents a technology used, while the two columns are the two deterrence strategies used by Asyr. On the horizontal axis, the ASW time shows how long it took for each technology to satisfy mission requirements. On the vertical axis, each technology has different values representing the number of investments made. For instance, in NATO Sonar Quality technology, it is possible to see that the time is uniformly spread, meaning that this technology does not make the mission faster, and that the technology level reaches values of 5 and 6, meaning that strong investments are needed. Looking at first at the passive deterrence strategy two technologies are the key: those increasing sonar range and those increasing the payload of the helicopters (i.e., number of sonobuoy fields available). This is in line with expectation for scenario types with no active engagement.

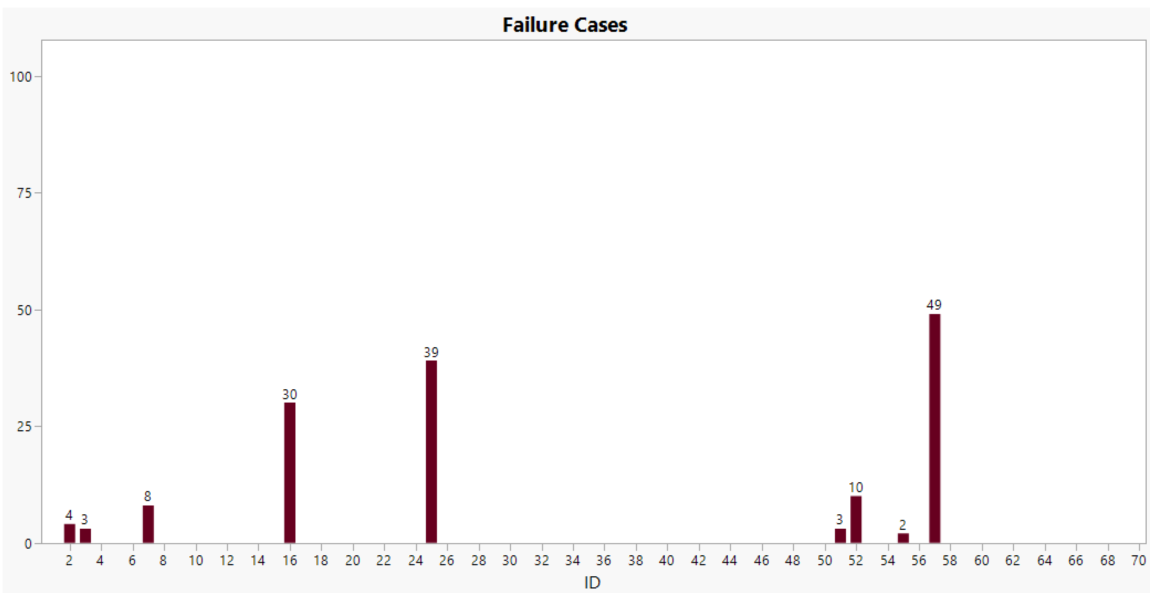


Figure 34: Scenarios in which the fleet still fails after the iterative algorithm (in red aggressive deterrence strategies, passive deterrence strategies are not present)

Regarding active deterrence strategies the issue is more complex. Looking again at Figure 35, a clear need for better detection mechanism emerges, with an increase in both sonar quality and range. Ships defensive systems need also to be upgraded mostly through an increased number in decoys availability and decoys' quality. Limited investments are needed in ships' hull strength. Defensive mechanisms for helicopters also show the need for upgrades with some investments needed in the number of flares available and the quality of those flares. Finally, regarding attacking technologies, as shown in Figure 35, heavy torpedoes are preferred as well as weaponizable helicopters that can attack underwater vessels. Factors like the number of torpedoes on the frigate, the helicopter endurance and the number of helicopters were not considered in need of upgrades by the algorithm.

Figure 35: Contour diagrams of the resulting technologies of interest after the algorithm runs.

7.4 Conclusion

In conclusion this experiment was successful in proving that this algorithm can indeed find rapidly several technology strategies that reduce the failure rate of the fleet in Vulnerable Scenario. This experiment demonstrated that all the deterrence strategy cases were resolved as well as the vast majority of the aggressive deterrence strategy ones.

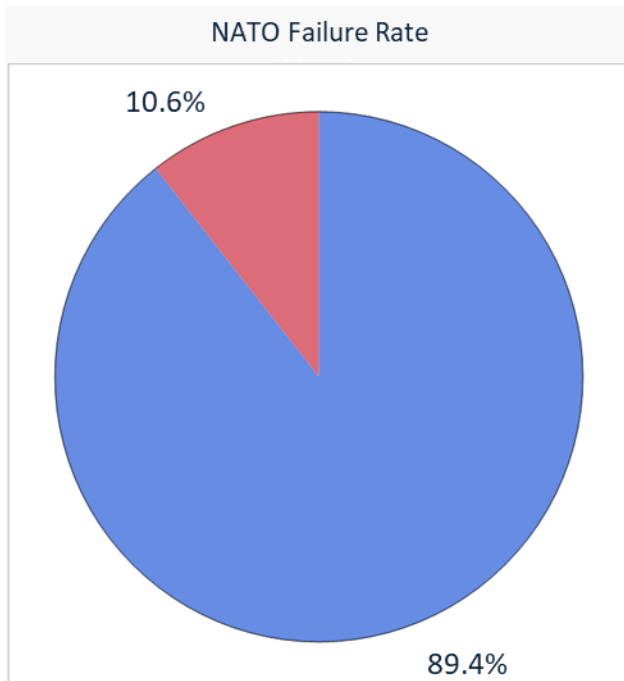


Figure 36: Failure rate of the evolved fleets in the Vulnerable Scenario. About 240 scenarios failed.

Figure 36 shows exactly the proportion of the scenarios that the algorithm was able to address. With a success rate of almost 90% the code was able to increase the technological level of about 2000 fleets to the point of success. For the remaining unsuccessful 10% Figure 37 provides an explanation of the different causes that caused failure. We see that about 73% of the cases failed due to the killing of at least one of the helicopters. Breaking down this 73%, 67% just helicopters, 4% frigate and helicopters and 2% helicopters and time. From this, it is possible to infer that the set of technologies of interest provided did not address the safety of the helicopters in some of the scenarios. Because this work focuses on S&T, other issues as tactical choices will not be considered.

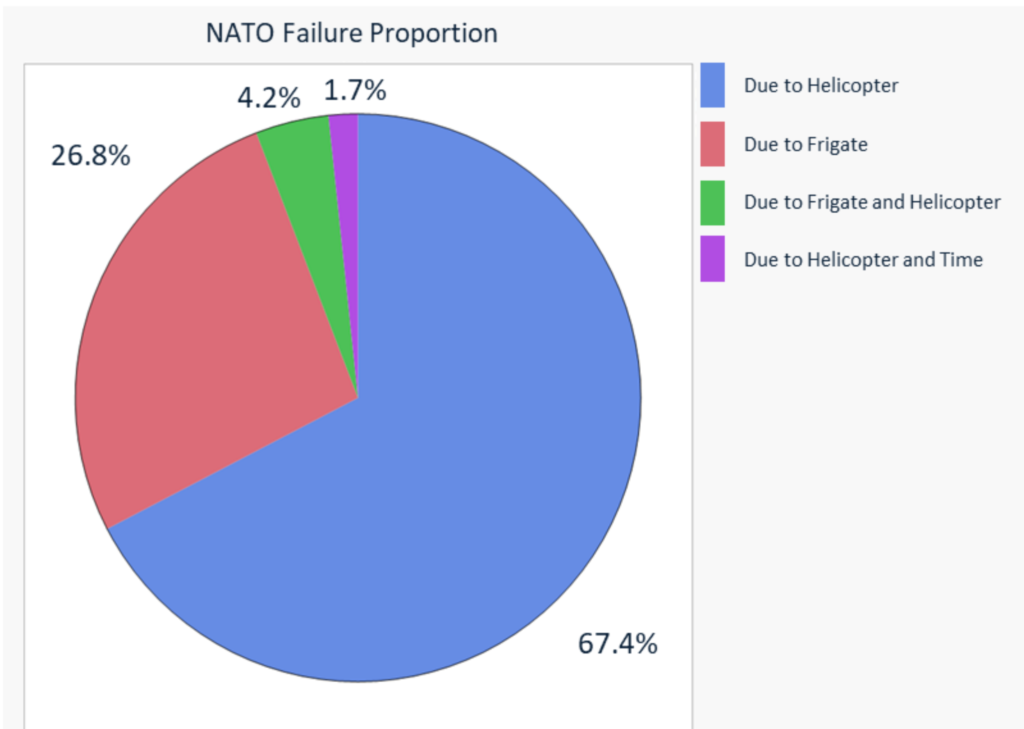


Figure 37: Proportion of the different failure causes.

Moving on to the iterative algorithm per se, it is interesting to check how far off the iteration cap was set up at 50. Figure 38 shows that most fleets are made successful in less than 10 iterations, with a median value for the whole experiment of 6 iterations. The maximum number of iterations to achieve success was 36. All the cases above 36 reached the cap value of 50 and failed, meaning that if time becomes a constrain the cap could be reduced to 20 or 25 iterations losing from 6% to 1% of successful fleets. If that strategy is adopted the total number of runs wasted for unsuccessful fleets will drop down from 12000 which is 240 cases times 50 runs, to 4800. The average number of iterations to make a fleet successful was 7.9, which means that the total number of runs for the successful fleets was about 16300. From all these considerations it appears that currently 42% of the computational time was wasted on unsuccessful cases. By reducing the iteration cap to 20 iteration the wasted computational time can almost halved to 22%, with a loss of only about 6% of successful fleets.

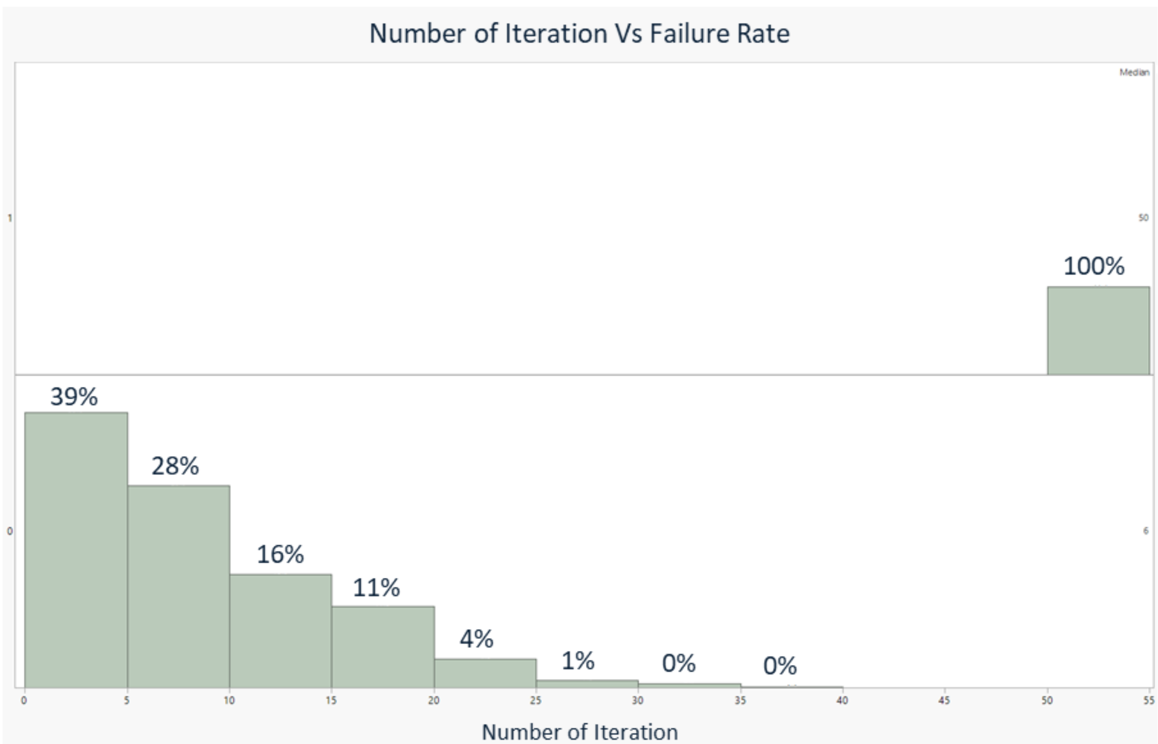


Figure 38: Number of algorithm iterations needed to achieve success compared to the success rate

To understand what are the causes that lead to failure it is possible to use a correlation map of the scenario variables, success indicators and fleet variables, reported in Figure 39. From this, it emerges that when the fleet failed due to the frigate being killed the algorithm was focusing on increasing the quality of decoys and the hull strength. Regarding scenarios in which helicopters were killed, there is a strong correlation with the increase in flares availability and in the quality of these flares. As it was shown in Figure 37, time by itself never caused the fleet to fail. From this analysis it is possible to infer that the algorithm was working in the right direction by increasing variables of technologies related to the specific problem. Nevertheless, even if the actions performed by the algorithm were correct, they were not enough to make the fleet successful.

Conjecture 6: Because the change of a technological parameter is driven univocally by specific values in the outputs, the order of technologies changed in the tree is not relevant.

While no proof was required for this conjecture, looking at the order of technologies shown in Figure 33 and at the results from Figure 35 our conjecture was correct. his was demonstrated and can be seen in the way works, but it was also confirmed by the fact that the technologies chosen were taken from independent levels in the algorithm.

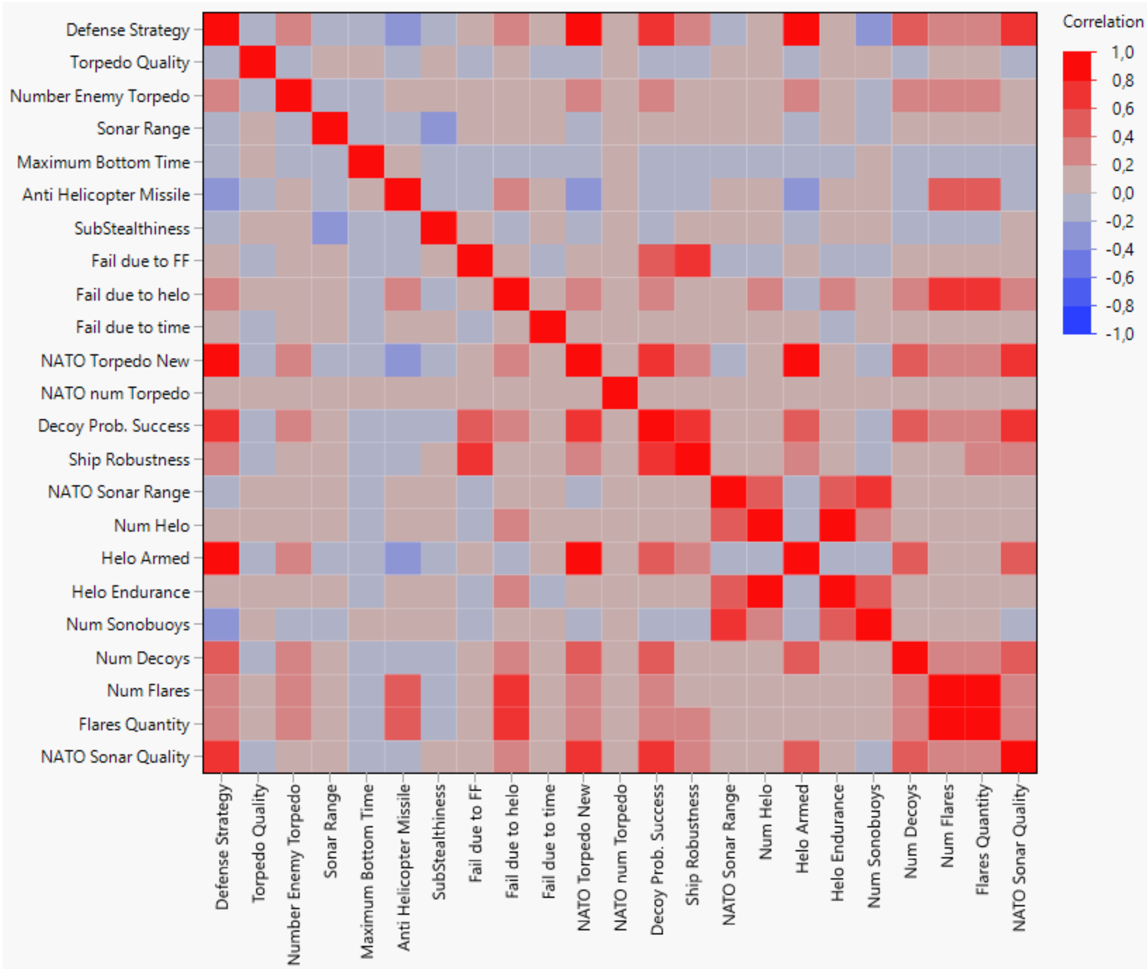


Figure 39: Correlation map of the scenario variables (row 1-7), success indicators (row 8-10) and fleet variables (row 11-23).

CHAPTER 8. Non-Positive Monotony of Technologies

The focus of this experiment is to verify Hypothesis 4 and Hypothesis 5 by checking that a set of technologies does not cause the opening of other Vulnerable Scenario and that the negative effects can be mitigated by investing in several technologies in parallel. To do so this experiment will rely on the results of the experiment in Chapter 6 (i.e., the Vulnerable Scenario for the ASW case) and on the set of technologies of interest highlighted in Chapter 7.

To review the statement of Hypothesis 5: If one technology has only a positive impact on the fleet, then its use will not lead to new Vulnerable Scenarios and non-Vulnerable Scenario will remain as such. In other words, if there are technologies that negatively affect the fleet, the impact is that it might open new Vulnerable Scenarios on top of those already found. The goal of the experiment is therefore to check if there are new Vulnerable Scenarios, or not, when the effects of technologies of interest are maximized. Hypothesis 4 will also be tested in this experiment as it states that that it is expected that when all technologies are maximized, they will be able to reduce the effect of few technologies negatively impacting the fleet.

8.1 Methodology

The experiment follows a simple methodology. First, fleets are generated by infusing one maximized technology into the basic fleet, to these is then added a fleet with all maximized technologies. By doing so, n+1 fleets are generated, where n is the number of technologies of interest. If there is a technology that has negative effects on the fleet this will appear by creating a new Vulnerable Scenario.

Once the fleets are created, they need to be tested in a set of scenarios. There are two options, they can either be tested on the full set of scenarios, or they can be tested only in the subset of non-vulnerable scenario. The advantage of taking the first option is that it is possible to use the same algorithm already developed to find the Vulnerable Scenario, achieving a more direct comparison with the results previously obtained. The disadvantage is that the number of cases to run is much higher. The second option gives a quicker way to see if there are new Vulnerable Scenario as the number of cases to run is lower. Given that the algorithm previously developed through PRIM cannot be used – as the set of data is different – a different approach must be used. A feasible way to ensure that the results are the same is to conduct an analysis on JMP to see how many cases failed per ID. If the results are the same compared to what was achieved by the Basic Fleet, then it could be inferred that there are no new Vulnerable Scenario. If the results are different, then further analysis will be needed to understand what went wrong and to identify which variable produces the negative effect.

8.2 Experiment

For running this experiment, it was decided to follow the first approach. Therefore, all the 7000 cases generated for the first experiment, the one described in Chapter 6, were run again n+1 times. In this case, in fact, there are no limitations on computational time, and we are just interested in evaluating the methodology, it therefore makes sense to go for the approach that provides the most comprehensive results. As discussed in Chapter 7 the total number of technologies to be tested was 12. To create the pool of fleets to be tested, each technology was maximized in each fleet, plus one additional fleet was added to test the complete set of technologies maximized. This created 13 different fleets and increased the number of cases to be tested to 98000. To ensure consistency with the previous experiment the random seed for each of the 7000 different scenarios was kept constant.

Table 9: Maximized Parameter per Fleet

Fleet	Modified Parameter
1	Basic Fleet
2	Sonar Range Maximized
3	Helicopter Number Maximized
4	Sonar Quality Maximized
5	Torpedo Number Maximized
6	Hull Strength Maximized
7	Helicopter Armed
8	Flight Endurance Maximized
9	Sonobuoy Fields Number Maximized
10	Decoys Quantity Maximized
11	Torpedo Decoy Quality Maximized
12	Flares Number Maximized
13	Best Flares Quality
14	All Variable Maximized

Regarding the technology variables, the values used can be found in Chapter 7, Table 7and Table 8. For this experiment, as the selection of the technology level in the fleet was binary, only the max and min values were used.

All the cases were prepared on JMP, where a csv file was arranged with the whole batch of cases to be run. As in all the other experiments described so far JANUS was used as Agent-Based modeler software. Therefore, after the csv file was created it was loaded on JANUS to run the experiment. The scenarios were outputted disregarding the success of the fleet – differently from the experiment run in Chapter 7. The success criteria used are the following: the mission should be completed within 30.000 time units and without any loss of any major assets. If these success criteria are not met the case is marked as failure to help PRIM in identifying scenario boxes. The output came in the form of another CSV file which was first uploaded to JMP, and then run through PRIM to identify the Vulnerable Scenarios. The former step was needed to find if negative impacting technologies were present, while the latter to see if the PRIM boxes changed or not. If there are changes in the PRIM box then, according to Hypothesis 5 some of the technologies will negatively affect the fleet.

Depending on if and how much the PRIM boxes change it can be the symptom of some variables negatively affecting the fleet. If that happens, the analysis on JMP becomes critical to identify those variables and if their effect can be mitigated by other variables. In Hypothesis 4 it was stated that such behavior is expected, however, to demonstrate this hypothesis there is no need to have changes in PRIM boxes. If Hypothesis 4 is verified, then the output from JANUS will show a much lower failure rate in the fleet with all maxed technologies than in the other 13 where only one technology is maxed.

8.3 Results

Looking at the result from JANUS only the 0.08% of the 98,000 simulations failed in achieving any outcome, which is well within the acceptable margins. The results from the PRIM analysis of the outcome dataset can be seen in Figure 40. The PRIM boxes that emerged are like the one in Chapter 6, but they do not overlap completely, the comparison can be seen more explicitly in Table 10. Further analysis of this result will be presented in the conclusion of this chapter after all the experimental results have been presented. Looking at the PRIM box parameters reported in Table 10 the PRIM run was successful, and it achieved comparable results of what was achieved in Chapter 6. The coverage is slightly higher, while the density is slightly lower. These results are in line with was achieved before, the small fluctuation can be attributed to the different number of cases tested in this experiment.

Table 10: Comparison of PRIM results between the two experiments

Variable	Chapter 6 Range	Current Range
Num. of enemy torpedoes	6 – 14	10 - 14
Sonar range	0 – 5	3 – 7
Max bottom time	2 – 6	2 – 5
Sub. stealthiness	7 – 8	6 – 8

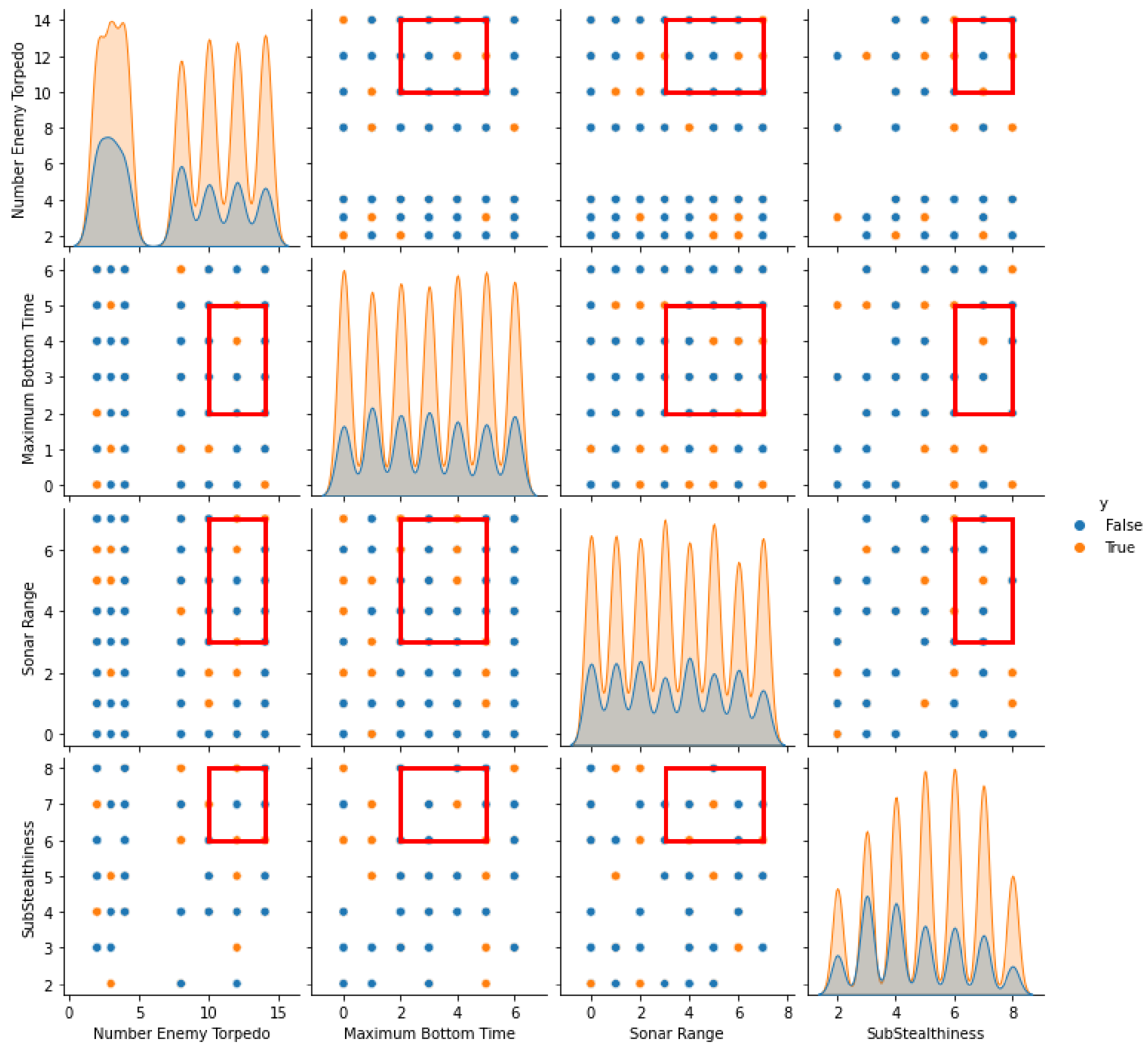


Figure 40: PRIM analysis of the result from experiment 3 (point 7), the red boxes represent the set of Vulnerable Scenario. In blue cases where the NATO fleet had success, in orange where it failed.

By running the same results on JMP it is possible to see how each technology contributes to the overall success of the fleet. The whole results are presented in Appendix E. but the subset in Figure 41 shows the key points. This figure demonstrates the failure rate. Recalling that fleet number 1 is the Basic Fleet while all the others have one (or all) technology maximized; it is expected that those technologies positively affecting the fleet will drive down the number of failed cases, while those technologies negatively affecting the fleet will increase that same number. From Figure 41 it emerges that there are 3 technologies shown in fleets 7, 8, 9 - all helicopters' variables: helicopter armed, helicopters' endurance and number of sonobuoys respectively - that in aggressive deterrence strategy affect the fleet. One technology is negatively impacting the fleet in the passive defense case. From these results it also emerges that there is one technology in fleet 5, sonar range, that is strongly positively affecting the fleet. The combination of all maxed technologies is also success with a strong reduction in failure rate. While Figure 41 is only an extract it coherently represents the results reported in Appendix E.

Table 11: Experiment 3 PRIM iteration results

Mean	0. 864796
Mass	0.1
Coverage	0.123001
Density	0. 864796
Restricted Dimensions	4

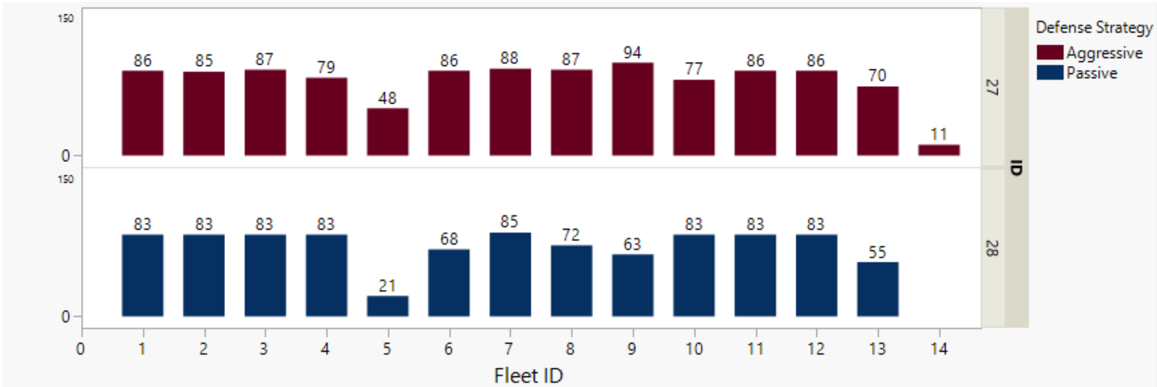


Figure 41: Extract of the results from experiment 3

8.4 Conclusion

Comparing the PRIM boxes obtained in this experiment with those from Chapter 6 we see that there are some differences but that those differences are not extensive. This is a symptom of the fact that the technologies can generate some different Vulnerable Scenarios, but that the effects are not strong. We see that these negative effects are fully mitigated by the combined use of maximized technologies, as demonstrated with fleet 14, proving that Hypothesis 4 was correct. The fact that there are technologies negatively impacting the studied fleet, and that this is also reflected in the PRIM boxes, also validates Hypothesis 5.

The negative effects of the different technologies are dependent on the deterrence strategy. The fleet composition and technology combination increase the risk of being shot by a submarine equipped with anti-helicopter missiles, this is shown by the results of fleet 8. In parallel, it should not come as a surprise the fact that in a scenario where helicopters do not get attacked the same variable will reduce the failure rate.

Finally, it must be mentioned that this method enables capturing which technologies are expected to have the biggest impact on the fleet. The reason results are different from what was obtained in Chapter 7 is that those technologies do not work in isolation. In this experiment the focus was to understand the impact of each technology in a binary way: positive or negative. As such, each technology was maxed only once. On the other hand, in Chapter 7 these technologies had to work together creating more benefits to the fleet. This conclusion is also supported by the fact that fleet 14 has a consistently higher success rate than fleet 5, proving that it is not just one technology driving success but that the key is the combined effort.

CHAPTER 9. Fleets Aggregation Strategies

While drafting the comprehensive methodology to answer the main research questions in Chapter 6, the need to be able to reduce the number of fleets studied arose. This concept is summarized in Hypothesis 6, which is the focus of this chapter. Hypothesis 6 states that between the two approaches drafted to minimize the number of fleets, maximized fleet approach and non-dominated fleets approach, the first one will produce fleets with a higher success rate.

This chapter is therefore investigating this hypothesis by comparing the two different sets of fleets studied. Moreover, it will also provide a demonstration for the final piece of the overall methodology: the criteria used to select the fleets to be tested in all the Vulnerable Scenario. It is important to recall at this point that during the overall methodology the number of fleets generated through the iterative algorithm was more than one per scenario. The total number was the number given by the number of Vulnerable Scenario times the number of repetitions. Testing all those fleets in all the Vulnerable Scenario (times the repetitions) would be prohibitive, therefore we were interested in finding a way to reduce the number of fleets. This chapter demonstrates why it was decided to study only those fleets generated by maximizing the value of each technology reached in each vulnerable scenario. This reduced the number of fleets tested to one per vulnerable scenario.

The reason this experiment was selected is that we wanted to study how relevant is time in the fleet selection process. With this objective in mind, the experiment was shaped in such a way that a comparison between the two different alternatives is conducted. This experiment is therefore needed not only to verify Hypothesis 6, but also, to ensure that satisfactory results can be achieved in a reasonable timeframe.

9.1 Introduction

The first alternative to be studied is a set of fleets created taking the maximum values of each technology from each evolved fleet. In other words, different evolved fleets are arranged by Vulnerable Scenario, the maximum value for each technology is taken among the fleets in the same Vulnerable Scenario. By doing so, from the original number of evolved fleets – which was equal to the number of repetitions originally chosen – the number of fleets is reduced to one per scenario.

The advantage of this first alternative is that the number of fleets to be tested is extremely limited, leading to shorter experiments. On the other hand, because the fleets studied are expected to have many, if not all, technologies to the maximum level, it is anticipated that they will be the most expensive ones to invest in. At this stage of the design process, no consideration about physical requirement was made yet but maximizing all technologies might lead also to prohibitive physical requirements.

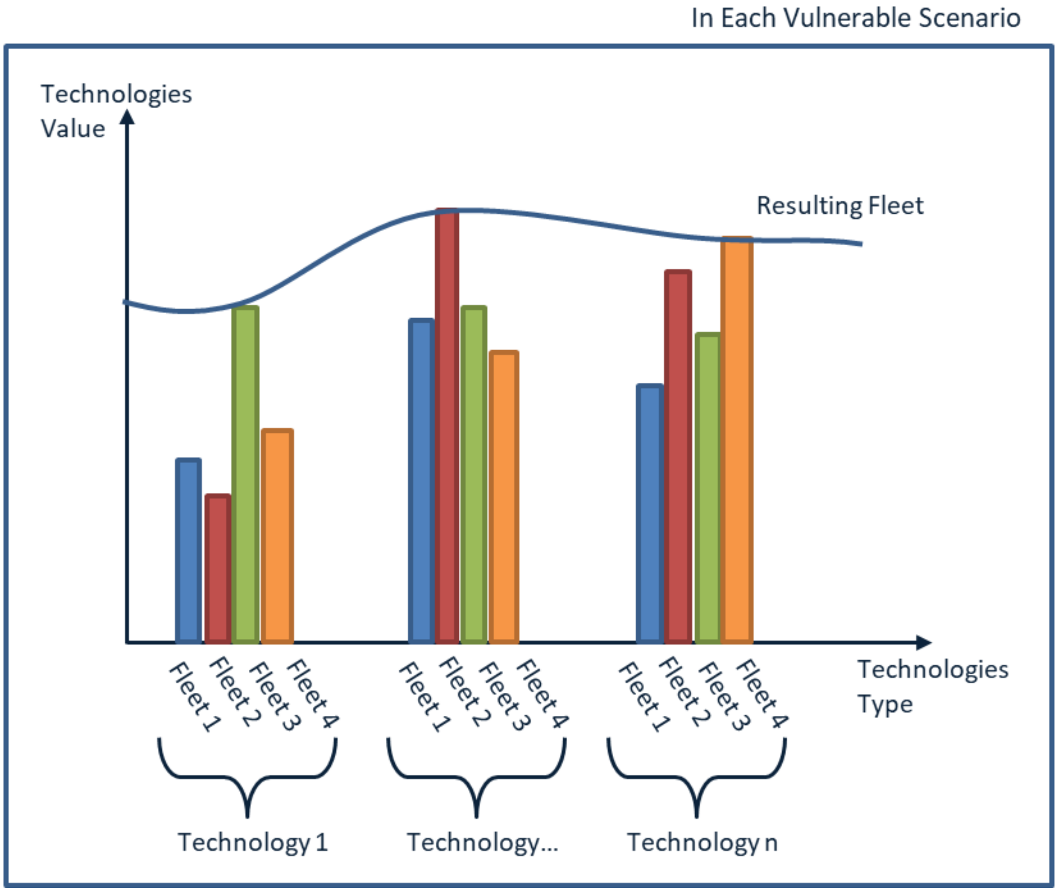


Figure 42: Representation of how fleets are generated in alternative one. The process is repeated in each Vulnerable Scenario

The second alternative is to study non-dominated fleets. In each Vulnerable Scenario, the non-dominated fleet, located on the multidimensional Pareto surface, are detected, and kept for further studies. The number of non-dominated fleets discovered depends on the number of technologies at hand - the higher the quality of technologies, the higher the number of non-dominated fleets. The main issue of using this approach compared to the first one is that it is necessary to test a much higher number of fleets compared to the previous one, about 7 times more as it will be shown later.

The benefit of this second alternative is that the studied fleets are not by definition also the most expensive one. In fact, since we are looking at non-dominated solutions, they might have a maximum value in one of the variables but not necessarily on another one. This leads to the comparison of fleets that have been successful in a Vulnerable Scenario, but that are not fully maximized across all technologies. The main negative aspect of this alternative is the high number of fleets that must be tested in all scenarios of interest.

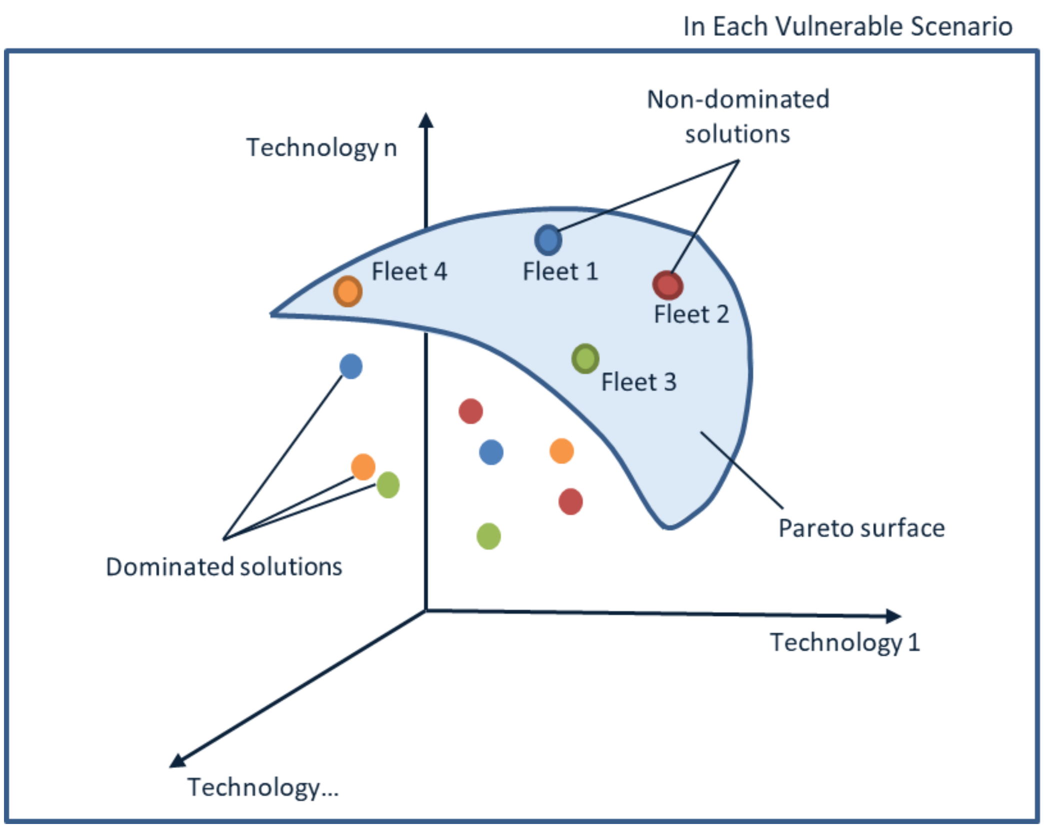


Figure 43: Representation of dominated and non-dominated fleets as generated in alternative 2.

9.2 Methodology

The methodology in this experiment is straightforward and can be divided into 4 steps:

1. Two datasets are created. These sets should be the byproduct of a previous study in which an optimization algorithm has been used to create Evolved Fleets.
    a. The first set is obtained by taking the maximum for each technology of interest in each Vulnerable Scenario, de facto creating new fleets.
    b. The second set is created by looking at the value reached by each technology of interest, and then using those to select only non-dominated fleets for further analysis.
2. Both datasets are then uploaded on the agent-based modeler, where the simulations are run.
3. The outputs of the simulations are analyzed to see how effective the obtained fleets were with the two different approaches.
4. If the fleets obtained through the first alternative have a larger success percentage than those generated via alternative 2, then Hypothesis 6 is verified and the experiment is considered a success.

If Hypothesis 6 is verified, it means that it is possible to reduce the number of simulations needed to achieve comparable, if not better results.

9.3 Experiment

Following the methodology laid down in the previous section, the experiment was set up. The data used for this experiment was derived from the results of the experiments run in Chapters 6 and 7. The Vulnerable Scenarios used are the 23 discovered ones in Chapter 6, while the Evolved Fleets tested are those created by the iterative algorithm used for the experiment in Chapter 7. Fleets' success criteria were also taken from Chapter 7, as such a friendly fleet is considered successful if it completes the mission of locating or neutralizing 2 submarines within 30.000 time units without any loss of any major asset, being them the frigate or one of the helicopters.

Regarding the first dataset, the data on evolved fleet from Chapter 7 was arranged per vulnerable scenario. From the 2.300 evolved fleets, 23 groups of 100 fleets each were made. 23 new fleets were created by taking the maximum value of each technology of interest among the ones from the 100 repetitions. These 23 fleets were then checked for repetitions, which once eliminated reduced the number of unique fleets to 10. These fleets are reported in Table 12.

Table 12: List of fleets tested in the first alternative

Fleet ID	New torpedo	Torpedoes **Number**	Decoy Probability	Hull Strength	Sonar Range	Helicopter Number	Helicopter Armed	Helicopter Endurance	Sonobuoy Number	Decoy Quantity	Helicopter Flares Number	Helicopter Flares Quality	Sonar Quality
1	1	6	10	10	10	2	1	6	3	10	2	0	6
2	1	6	10	10	11	2	1	6	3	10	4	1	6
3	1	6	9	5	11	2	1	6	3	8	4	1	5
4	1	6	10	9	10	2	1	6	3	8	4	1	6
5	1	6	10	7	11	2	1	6	3	9	4	1	5
6	1	6	10	10	11	2	1	6	3	8	4	1	6
7	1	6	10	10	9	2	1	6	3	10	4	1	6
8	1	6	10	10	11	2	1	6	3	10	2	0	5
9	1	6	10	10	11	2	1	6	3	9	2	0	6
10	1	6	10	10	11	2	1	6	3	10	2	0	6

For the second dataset, the starting point was again the set of evolved fleets from Chapter 7. As in the previous case, 2.300 fleets were arranged by Vulnerable Scenario creating 23 pools of 100 fleets each. At this point, the fleets were filtered to remove duplicates and then they were sorted to separate dominated from non-dominated solutions. In this process, the focus was to maximize each of the 13 variables. Therefore, when looking at non-dominated solutions the interest was in those fleets that had at least one maximum values across the set of technologies. Since the duplicate check was done within each individual vulnerable scenario, it had to be redone once the fleets were aggregated. This last step reduced the total number of fleets generated to 58. The full list of fleets is reported in Table 13.

Table 13: List of fleets tested in the second alternative

Fleet ID	New torpedo	Torpedoes Number	Decoy Probability	Hull Strength	Sonar Range	Helicopter Number	Helicopter Armed	Helicopter Endurance	Sonobuoy Number	Decoy Quantity	Helicopter Flares Number	Helicopter Flares Quality	Sonar Quality
1	1	6	10	10	5	1	1	3	1	10	2	0	6
2	1	6	5	1	10	2	1	6	2	8	2	0	6
3	1	6	10	10	4	1	1	3	1	4	2	0	5
4	1	6	0	0	9	2	1	6	3	2	2	0	1
5	1	6	9	10	8	1	0	3	1	2	2	0	6
6	0	6	0	0	10	2	0	6	3	2	2	0	0
7	1	6	10	10	0	1	1	3	2	10	2	0	3
8	1	6	9	10	8	1	1	3	1	2	2	0	6
9	0	6	0	0	11	2	0	6	3	2	2	0	0
10	1	6	10	10	8	1	1	4	2	4	2	0	6
11	1	6	10	10	8	1	1	3	1	2	2	0	6
12	1	6	10	10	0	1	1	3	2	10	2	0	4
13	1	6	5	0	11	2	1	6	3	8	2	0	2
14	1	6	10	10	8	1	1	3	3	4	4	1	6
15	0	6	10	10	0	1	1	3	2	10	2	0	0
16	1	6	10	10	11	2	0	1	3	10	4	1	5

17	1	6	10	10	8	2	0	6	3	10	4	1	6
18	1	6	5	4	11	2	0	6	3	8	4	1	5
19	1	6	10	10	5	1	1	3	2	10	4	1	5
20	1	6	10	10	0	1	0	3	2	10	4	1	0
21	1	6	9	10	11	2	0	6	2	2	4	1	6
22	1	6	6	3	11	2	1	6	3	4	4	1	0
23	1	6	10	10	0	1	1	3	2	10	3	0	5
24	1	6	9	10	8	1	0	3	2	4	4	1	6
25	1	6	6	4	10	2	1	6	3	2	4	1	1
26	1	6	5	4	8	2	1	6	3	8	4	1	5
27	1	6	10	9	1	1	0	3	2	4	4	1	1
28	1	6	6	3	8	1	0	3	3	5	4	1	6
29	1	6	10	5	8	2	1	6	3	4	4	1	1
30	1	6	6	5	4	1	1	3	2	3	2	0	5
31	1	6	5	0	4	1	1	3	2	9	4	1	0
32	1	6	8	5	8	2	1	6	2	2	4	1	6
33	0	6	10	10	0	1	0	3	1	10	4	1	0
34	1	6	6	2	9	2	0	6	3	7	4	1	5
35	1	6	6	4	11	2	0	6	3	4	4	1	5
36	1	6	5	0	8	2	0	6	2	10	4	1	6
37	1	6	10	10	1	1	0	3	2	10	4	1	0
38	1	6	5	0	8	2	1	6	3	10	4	1	3
39	1	6	9	10	12	2	0	6	2	5	4	1	6
40	1	6	10	10	2	1	1	3	2	4	2	0	5
41	1	6	9	5	8	2	1	6	3	7	4	1	5
42	1	6	1	0	11	2	1	6	3	3	2	0	1
43	1	6	10	5	8	1	1	5	3	10	4	1	5
44	1	6	10	8	2	1	1	3	2	10	4	1	3
45	1	6	6	6	8	2	0	6	2	10	4	1	6
46	1	6	6	10	8	1	0	3	1	5	4	1	6
47	1	6	10	10	5	1	1	3	2	10	4	1	3
48	1	6	1	0	8	2	1	6	3	5	4	1	0
49	0	6	5	0	2	1	0	3	2	10	4	1	0
50	1	6	10	10	0	1	1	3	1	3	2	0	5
51	1	6	5	0	8	2	1	6	3	10	4	1	0
52	1	6	10	10	0	1	0	3	1	2	4	1	0
53	1	6	10	5	11	2	1	6	3	3	2	0	5
54	1	6	10	10	5	1	1	3	2	10	2	0	0
55	1	6	5	0	11	2	1	6	3	9	2	0	2
56	1	6	10	10	0	1	0	3	1	2	2	0	4
57	1	6	10	10	0	1	1	3	2	2	2	0	3
58	1	6	8	5	8	2	1	6	3	10	2	0	1

Using the two fleet sets defined before, two new csv files were created. In each of these, each of the fleet sets was paired with all the Vulnerable Scenario and their repetitions. As it was done in previous experiments to ensure consistency, the random seeds for the Vulnerable Scenario were kept constant at what was generated in the first experiment in Chapter 6. In the end, the first dataset consisted of 23.000 cases and the second one of 133.400 cases.

In two separate instances the files were uploaded on the agent-based modeler JANUS and the different runs were conducted. Once the simulation was over, the results were then transferred to JMP to analyze the results.

9.4 Results and Conclusions

From the results of the experiment, it is possible to verify Hypothesis 6. In both datasets, the percentage of fleets failing to deliver outputs was below 5%, which was within acceptable ranges. By processing the fleets on JMP, it is possible to calculate the failure rate of each dataset. In alternative 1, was slightly below 20%, while in alternative 2 the failure rate was almost 40%. Figure 44 shows the comparison between the two alternatives.

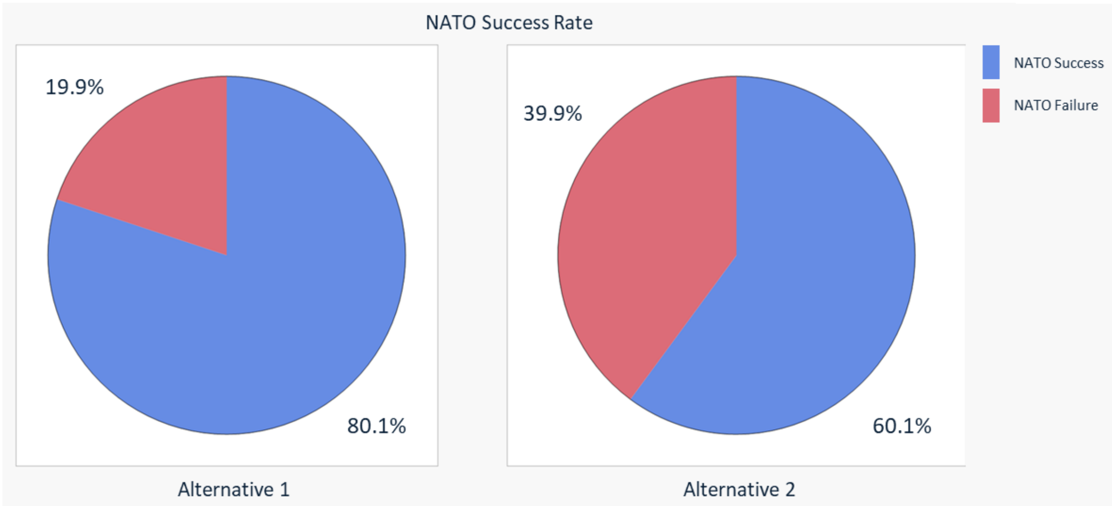


Figure 44: Success rate for the two tested datasets.

From the results it is clear that the first alternative has a much higher success rate compared to the second one. The higher success rate, with the information on the previous chapter regarding Hypotheses 4 and 5 hence verify Hypothesis 6. Although, it is possible to think that the higher success rate was driven by the stochasticity of the process this is not the case. In fact, by verifying both Hypotheses 4 and 5 it was established that the technologies were either positive monotone, or that their combined effect is. Looking now at how the two alternatives were generated, it is possible to observe that the average technological level of the fleets in the first alternative is higher than that of the second alternative. Combining this information with the one regarding the monotony of the technologies involved, means that there is a positive trend that links technology level and success rate. This means that results obtained can be generalized and that the first alternative produces fleets with higher success rates. Moreover, it is important to mention that these results were achieved in about 20% of the time compared to alternative 2. All of this increases the preference of using alternative 1 compared to alternative 2 when generating fleets.

This last experiment concludes the verification part of the 6 hypotheses needed for the overall methodology. All the elements that were verified throughout these past 4 chapters will now be aggregated together to demonstrate the overarching hypothesis and to satisfy the research statement.

CHAPTER 10. Full Methodology Demonstration

This Chapter is dedicated to proving the methodology depicted in Chapter 5 as whole. In the previous 4 chapters a specific ASW use case has been used to validate the different hypotheses of this thesis. In this chapter the methodology will be used on a different use case, and it will rely on the elements previously demonstrated to find the Robust Fleets.

In this chapter the goal is to demonstrate the Research Objective of developing a methodology to support trades-offs among naval assets and technologies, to assist investments on new maritime technologies in future threat scenarios. Following the methodology depicted in Chapter 5, the experiment structure has been divided into 3 parts. The first part is focused on identifying Vulnerable Scenarios, the second part on finding Evolved Fleets, and the third part on comparing those fleets to see which one is the most robust one.

10.1 Experiment Background

Before starting the experiment, it is important to provide some background on what the experiment is about and on the scenario that will be used. The goal of the experiment is to reproduce the same outcomes of a CBA, and it will do so by testing one Basic Fleet in multiple scenarios. All these scenarios are created by instantiating the uncertainties on the enemy side, nevertheless the main mission of the scenario will be constant. The reason why this goal was chosen is that it allows us to prove that a methodology that finds preliminary, and yet useful, results is possible. Moreover, by providing the same results of a CBA in a quantitative way it is possible to see the role of this experiment in the grand scheme of the investment problem.

The experiment will be successful if a set of Vulnerable Scenario and a set of Robust Fleets are produced at the end of Part 1 and 3, respectively. While it is not possible to validate the quality of the solutions, this does not compromise the validity of the experiment as the goal is to demonstrate the methodology, not the specific use case.

10.1.1 Scenario

The scenario is located in the fictitious archipelago of Relas, here there are 4 countries: Relis – considered a NATO allied, Asyr – considered a NATO adversary, and Prucy and Turim both neutral but respectively aligned with Relis and Asyr. Relis and Asyr share the biggest island of the archipelago while Turim is located on the north and Prucy on the west. Figure 45 shows the archipelago composition and the respective territorial waters.

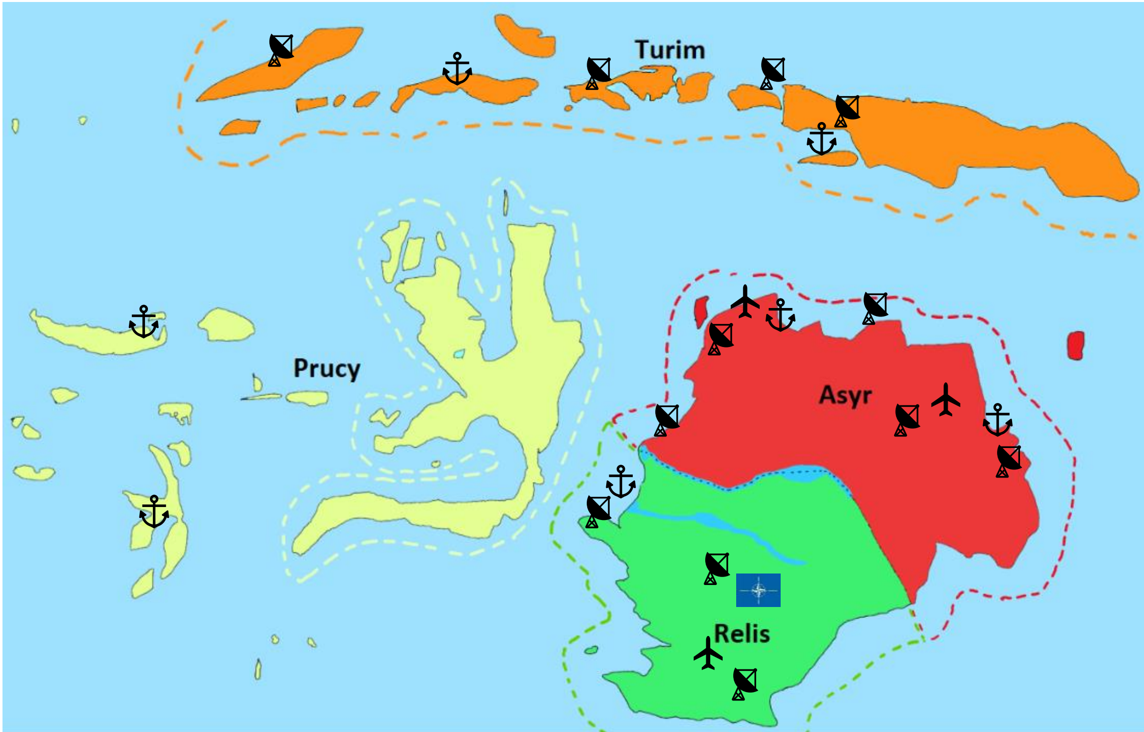


Figure 45: Relas Archipelago political map, dotted lines represent territorial waters.

After a snap attack by Asyr on Relis, Relis asks for support from its NATO allies which deploy a naval task force. The task force will have to reach the area in front of Relis' only seaport and evacuate civilians. During the first attack, Turim, in agreement with Asyr, decides to extend its territorial waters in front of Prucy, forbidding the passage of any military ships, de facto blocking the north channel. NATO forces will have to approach the port of Relis from the south. While getting closer to the island they will be exposed to Asyr's air force and to Asyr's anti-ship coastal defenses. NATO intelligence has also discovered several new coastal defense trucks which were not known, meaning that NATO will have to face threats coming from multiple axis, which will not always be known.

The NATO fleet goals are to arrive in front of the shores of Relis to evacuate a large group of civilians from the only seaport Relis has, without losing any ship during the approach and the evacuation. Asyr's goal is to stop NATO from achieving its goal and to do so it will try 3 different tactics, and it will deploy several bombers and semi-movable missiles trucks. Asyr will try to conceal the trucks to gain advantage against the NATO fleet, as such, the position of these trucks changes from scenario to scenario. Moreover, Asyr will not always shoot at maximum range, but it will try to fire at closer range to reduce NATO's reaction time and to maximize the probability of kill of its missiles.

The reasons why this scenario was chosen are multiple. First, it allows for multiple tasks inside one single mission, which is good as it might be desired by the decision makers that assets should be able to deal with different tasks in a meaningful way. Secondly, this scenario involves several types of assets working in a collaborative way. It was interesting to test how having multiple assets behaving independently and yet in a coordinated fashion can bring benefits to the investment strategies. Finally, this scenario was chosen because it allows us to test a probable future mission that will see a fleet of allied vessels engaging adversary forces in an archipelago heavily defended.

10.1.2 XLRM Framework

As already mentioned in Chapter 6, the XLRM framework can be used to provide clarity on which are the different moving parts involved in the experiment. The X stands for uncertainty, which are the different variables and ranges on the Asyr side, reported in Table 14. The policy levers (L) are all the variables on the NATO fleet side, which will be used to identify the evolved fleet first in part 2 and the robust ones in part 3. Policy Levers can be found in Table 15 with their starting values and ranges. Relationships (R) are managed by the agent-based modeler JANUS. As it was done for other experiments, JANUS will be used also here to perform all the simulations. Metrics (M) are those outputs that can be used to track the progress and the evolution of the fleet – namely if it is successful or not. In this experiment it was chosen to look at how many of the defense stations are destroyed and at how many ships are lost in the process.

10.1.3 Assumptions

Due to the complexity of the scenario, there are a wide set of assumptions that must be considered. The whole list is reported in Appendix F., here only a small portion will be reported to provide the reader with context and with a better understanding of the problem at hand.

Regarding Asyr, it has one main control center which is located in the North part of the island. This center is in Asyr's main airport, and whenever its radar detects the incoming ships, it can deploy up to 20 aircraft to engage them. Asyr's defense stations are randomly distributed across the main island where Relis is located and the main island of Prucy, where Asyr forces have managed to deploy some troops. These defense stations have a variable percentage of being already detected when the fleet approaches the area. Each station is protected by up to 10 Man-portable air-defense systems (MANPADS) to make sure that the stations can resist some level of aggression from NATO. Asyr has 3 main ways of deterring against the approaching NATO fleet: Persuasive, Saturation and Saturation All-Out. The description of these strategies and the reasoning why they were chosen will be discussed in paragraph 10.1.6.

NATO has a small fleet of 10 ships – 7 frigates, 1 High Value Unit (HVU), 1 destroyer, 1 minesweeper, which will be approaching the archipelago from the South. The number of ships is constant and cannot be changed, the HVU includes a group of support ships and the aircraft carrier which acts together and for modelling reasons was aggregated as one. The ships move in a coordinated formation which changes depending on the relative position of the fleet to the destination. The formations were chosen following the recommendation of an expert in the field, and as such they were not used as variables in the study. All the modified technologies are applied to all the ships and not just to one, this specific hypothesis was picked as the focus of the work is to look at Science and Technology products which could benefit the whole fleet.

10.1.4 Input

Table 14 and Table 15 show the relevant input variables used by both Asyr and NATO, respectively. These are also the Uncertainties and the Policy Levers of the XLRM framework. In Part 1 of the experiment, a DoE is created to test 150 different scenarios. Only the inputs from Asyr's pool are used in this DoE. The value of each variable is taken within the range reported in Table 14. NATO's variables are kept constant to the minimum level, or to a 'false' if the variable is binary. In Part 2, the situation is inverted to the values found in the Vulnerable Scenario, while NATO's variables are free to change according to the iterative algorithm. In Part 3, all the input variables are kept constant to what was found in the previous two parts as Vulnerable Scenario and evolved strategies.

All these variables were taken with a common interest: finding those parameters that could affect tor support the most the incoming fleet on Asyr's and NATO's side, respectively.

With this rational in mind, it was decided to look at what is critical in a mission involving semi-movable defense stations, and of what could complement those defenses. Moreover, it is of interest to see when there are priority targets involved as that shifts the focus of the attackers making them more prone to adopt risky behaviors. The selection of variables chosen following these criteria are reported in the following tables, divided per country.

Table 14: Asyr Variables

Variable	*Range*	*Notes*
Priority Targets	False True	When this is true the NATO fleet will try to hit the specific location of the island to disrupt Asyr's strategy.
Number of Planes	0 – 20	Number of planes deployed to stop the incoming NATO fleet.
Missile per Plane	1 – 4	Each plane has a variable number of air-to-ground missiles.
Number of Defense Stations (DS)	0 – 10	These are semi movable Bal-E trucks that are distributed around Asyr and Prucy. Each has a certain number of missile racks.
Number of Waves Each DS can Fire	1 – 3	The number of missiles racks each DS has.
Number of Missiles per Wave	4 – 10	The number of missiles present in each missile rack.
Number of MANPADS	0 – 10	The number of Man-Portable Air-Defense System around each DS.
Deterrence strategy	Saturation Saturation – All out Persuasive	The deterrence strategy employed by Asyr. Both Saturation strategies are coordinated attacks; in the Persuasive strategy attacks are carried by each unit independently from others.
Missile Quality	Low Medium High	Missile quality represents the level of complexity of weapons employed by Asyr.

Defense Station Cover Level	0 – 10	The probability that DS will be visible to the incoming NATO fleet. If 0 all DS are immediately visible, if 10 they are all hidden.

Table 15: NATO Fleet Variables

Variable - Asset (if present)		*Range*		*Notes*
Sonar Range	Frigate	20.000 – 25.000		The sonar range is given in increment of different measures depending on the ship. Values are in meters, and they are the min and max possible. Destroyers have two radars, one conventional and one for early warning which can only pass information on the presence of the enemy, but not its characteristics.
	Destroyer	25.000 – 32.000		
		50.000 – 62.000		
	HVU	17.000 – 23.000		
HVP Ammunition		False	True	This ammunition doubles the probability of kill of point defense systems against incoming threats.
New Short-Range Missile		False	True	Increase the probability of kill by 20%
New Long-Range Missile		False	True	Increase the probability of kill by 20%
New TLAM		False	True	Increase the probability of kill by 20%
Naval Guns Number	Frigate	2 – 5		Each ship has several point defense systems that are used to defend the ship against incoming threats.
	Destroyer			
	HVU	3 – 6		
VLS Blocks	Frigate	2 – 6		Each ship has several VLS blocks, each of which has 8 missiles tubes. In each tube there could be either one long-range missile, one TLAM or 4 short-range missiles.
	Destroyer	5 – 9		
	HVU	4 – 8		

VLS Fire Rate	60 – 25	The fire rate of the VLS can be increased to launch more missiles in less time. In this case the numbers represent the interval in time units between two launches.
Ships Robustness	0 – 14	The ship's robustness represents how resilient the ship is against hits. The probability of kill of an incoming missile gets reduces by 3% every tick.

10.1.5 Output

In Part 1, the only simulation outputs that are necessary to the identification of Vulnerable Scenario are those related to the success of the mission – how many and which NATO assets are destroyed, and how many of Asyr's defense stations are destroyed. In Part 2, outputs from Part 1 are used to increase the performance of the NATO fleet and to drive the iterative algorithm. Part 3 focuses again on the same outputs used in Part 1 as it is tasked to identify which of the fleet is the most robust one. These outputs were selected after consultation with military experts, and after reviewing what are the key variables driving the decision-making process. Table 16 shows all the outputs of the simulation and their explanation.

Table 16: Output parameters of the simulation and explanation

Parameter	*Notes*
Total Run Time	This is the total running time of each case.
seed	The random seed used in the case, stored for repeatability.
Number Ally Killed	Number of major units destroyed by Asyr, if bigger than 0 the case is a failure.
Number HVU Killed	Number of HVU destroyed by Asyr, if bigger than 0 the case is a failure.
Number Destroyers Killed	Number of destroyers destroyed by Asyr, if bigger than 0 the case is a failure.
Number Frigates Killed	Number of frigates by Asyr, if bigger than 0 the case is a failure.
Number Minesweepers Killed	Number of minesweepers by Asyr, if bigger than 0 the case is a failure.
Total Short-Range Interceptors Available	Number of short-range interceptors available by the whole fleet.
Total Short-Range Interceptors Launched	Number of short-range interceptors launched by the whole fleet.
Total Long-Range Interceptors Available	Number of long-range interceptors available by the whole fleet. Only the destroyer and the HVU have them.
Total Short-Range Interceptors Launched	Number of long-range interceptors launched by the whole fleet.
Total TLAM Available	Number of TLAM available by the whole fleet.
Total TLAM Launched	Number of TLAM launched by the whole fleet.
Total TLAM Intercepted	Number of launched TLAMs intercepted by Asyr’s defenses.
Bad Configuration	Used to mark optimized configurations that still fail the mission.

Number of Defense Stations Destroyed	The number of defense stations destroyed by NATO. If this number is lower than the number of visible defense stations, then the mission is a failure.
Primary Target Hit	A parameter to verify that if there are primary targets on Asyr side those are hit. This is a bonus objective and even if not achieved the mission will not stop.In Part 2 it will be used to further expand NATO's fleet capabilities.
Kh-35 Launched	The number of missiles launched by the defense stations.
Kh-35 Intercepted	The number of missiles intercepted by the NATO fleet.
Who Shot First	A parameter to verify who was the first shooter between Asyr and NATO.
Hidden Defense Stations	The number of defense stations that are not visible to the NATO fleet at the beginning of the simulation.
Visible Defense Stations	The number of defense stations that are visible to the NATO fleet at the beginning of the simulation.
Scenario ID	ID of the scenario, stored for traceability.

10.1.6 Assets Behavior

In this simulated conflict the task of the NATO fleet is to reach a certain area to perform its mission. Asyr's goal is to stop NATO from reaching its destination location. To do so it will employ different strategies depending on the scenario.

10.1.6.1 <u>NATO</u>

Since NATO might be aware of some of the defense stations Asyr has deployed, the fleet will enter the area in close formation with two semi circles of frigates, shown on the left of Figure 46, shielding the HVU and the destroyer on the top left of the formation to provide additional AAW shielding. The destroyer is the key asset against aerial threat as it employs an early warning radar that can help identify threats at long distance. The fleet will maintain this formation until it enters the channel between Prucy and Relis, at that point the frigates will rearrange to provide support in case of attacks from both sides of the channels, shown on the right of Figure 46. In the final part of the approach, the frigates will change again creating a secure channel between the two islands, as shown in Figure 47. The destroyer will be north of the HVU to provide additional AAW support on the side that will most likely see an attack incoming.

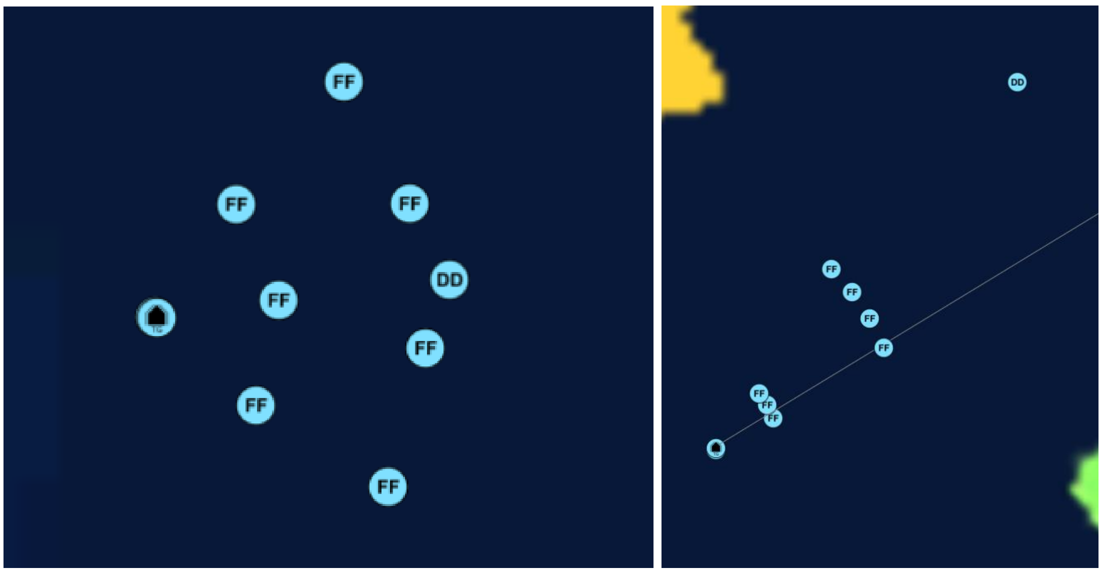


Figure 46: Formations of the NATO fleet at the entrance of the theater (left), and at the entrance of the channel (right)

The RoE for the fleet are to engage threats when they appear. This means that when a defense station shows up on the radar of the fleet one of the ships will fire to destroy it. Whenever there is a missile incoming or a threat enters in the range of the fleet, the ship which has identified it sends a message to the HVU. This can either task the same ship or another ship to act and fire appropriate missiles or countermeasures. This protocol is respected unless the incoming missile enters in proximity of a ship, at this point the ship uses its close in weapon system independently from what other ships are doing about that missile.

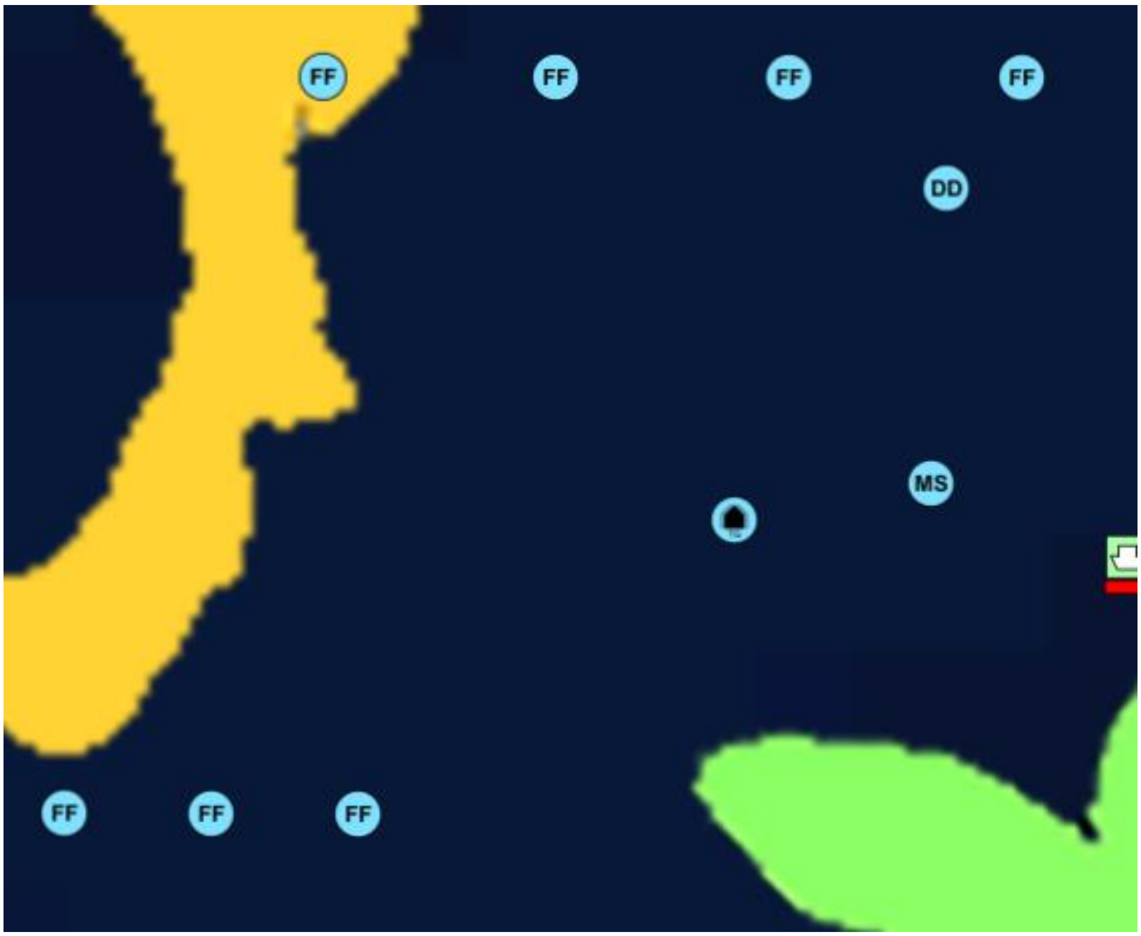


Figure 47: NATO Fleet position in defense of the HVU while performing operations

10.1.6.2 Asyr

As mentioned, Asyr can employ 3 different deterrence strategies: Persuasive, Saturation and Saturation – All Out. These strategies were selected as they represent well the several types of defenses that the country could set up, they vary in complexity, effectiveness, coordination needed and weak points. In all strategies, Asyr will distribute a series of semi-movable coastal defense systems across its territory and the Eastern side of the main island of Prucy. This operation is carried out before NATO's fleet arrival. Each truck has a variable probability of being hidden, if that is the case the truck will stay hidden – and therefore is not targetable by NATO forces. Once the NATO fleet approaches the archipelago one of the three strategies is employed:

- Persuasive: In the persuasive strategy, Asyr is trying to discourage NATO forces from approaching the archipelago. To do this, each defense station will fire their missiles whenever the fleet enters their range. Even if the defense distribution changes, in every simulation the rule is always that the closer to the destination point the more defense stations there are. This means that at a certain point multiple defense stations will fire from different axes. When the fleet enters Asyr's Air Force AoR, aircraft will be deployed from the airport in the north and will try to intercept the incoming fleet near Relis' port. In this strategy, there is no real coordination among Asyr assets. It represents a situation in which commanders of different units are isolated and cannot really talk.
- Saturation: In the saturation strategies all the assets are coordinated by the C2 center (which is also a priority target). Each defense station sends a message when the NATO fleet enters its firing range. When the fleet enters the Air Force AoR, aircraft are deployed and put on a loitering pattern around their final position. When the fleet is inside the ranges of all defense stations the C2 center orders to fire. Missiles are launched at the same time from all Asyr's platforms. Depending on where the

defense station is located, in the first wave, it will aim either at the destroyer or at the HVU. The idea is to do as much damage possible, by destroying the HVU, and to cripple NATO AAW capabilities, by sinking the destroyer. The targets for the second and eventually third wave are distributed among what is left of the NATO fleet. If the C2 center is destroyed, then each asset behaves like in the Persuasive strategy.

- Saturation – All Out: This strategy emulates the Saturation one with one caveat: missiles are not all launched at the same time, but they are launched in an ordered fashion so that they all arrive at the same time.

10.2 Part 1: Finding Vulnerable Scenarios

In the first part of the experiment, the goal is to find the Vulnerable Scenarios. The experiment can be summarized in the following points:

1. Because the focus of this part is the scenarios, the DoE is created considering only variations on Asyr's variables found in Table 14. NATO fleet's parameters in Table 15 are constant at the minimum level to represent the state-of-the-art. Moreover, this provides JANUS a complete set of inputs needed to create the scenario. Among the different possible DoEs, a space filling Latin Hypercube Design is used to capture the possible range of effects that scenario's variables have. The DoE is generated in JMP and then transferred to JANUS via a csv file.
2. Once the DoE is created and uploaded into JANUS, the simulation is run. The model runs following the assets' behavior described above to evaluate the performance of the fleet. Given that in the DoE 150 unique cases were created and

that following a bootstrapping analysis the number of repetitions was 80, JANUS had to run 12.000 scenarios.

3. Once the simulation is done, outputs are transferred from JANUS to JMP where the data is reformatted to perform the PRIM analysis.
4. After the PRIM algorithm is run Vulnerable Scenarios emerge as ranges in Asyr's input variables. These are used to update the information in JMP where the Vulnerable Scenario are displayed.

10.2.1 Results of the simulation

From the results displayed in Figure 48 it is possible to see how all 3 deterrence strategies offered a challenge to the fleet, with most cases failing to reach success. In general, Figure 43 depicts a grim picture for the NATO fleet, proving the need for upgrades to be successful in the tested scenarios. Among the 3 strategies, the Saturation All-Out deterrence strategy was less of a threat compared to the other two though. This should be attributed to the way incoming threats are countered in the simulation. In fact, whenever a missile or a bullet hits an object there is a minimum distance that triggers that detonation. This means that if incoming missiles are too close to each other, multiple missiles are destroyed by the same detonation, this is clearly one of the major limitations that was discovered while using JANUS. It is possible to reduce the explosion distance to below a certain threshold, but that has shown itself to cause issues in the code and therefore it was avoided.

In the correlation heatmap in Figure 49, it is shown that there is no prevalence of any specific factor in NATO's fleet failure. This is demonstrated by the fact that the correlation between any of Asyr's variables and the variable measuring NATO's success is not strong, never exceeding ±0.15. While this is expected for the correlation among Asyr's variables as they are correlated only by the DoE – which was created using a uniform distribution on purpose – this was not the case for the NATO failure variable. The result is therefore surprising, and it demonstrates that all the variables considered for Asyr were equally relevant and needed in the description of the Vulnerable Scenarios.

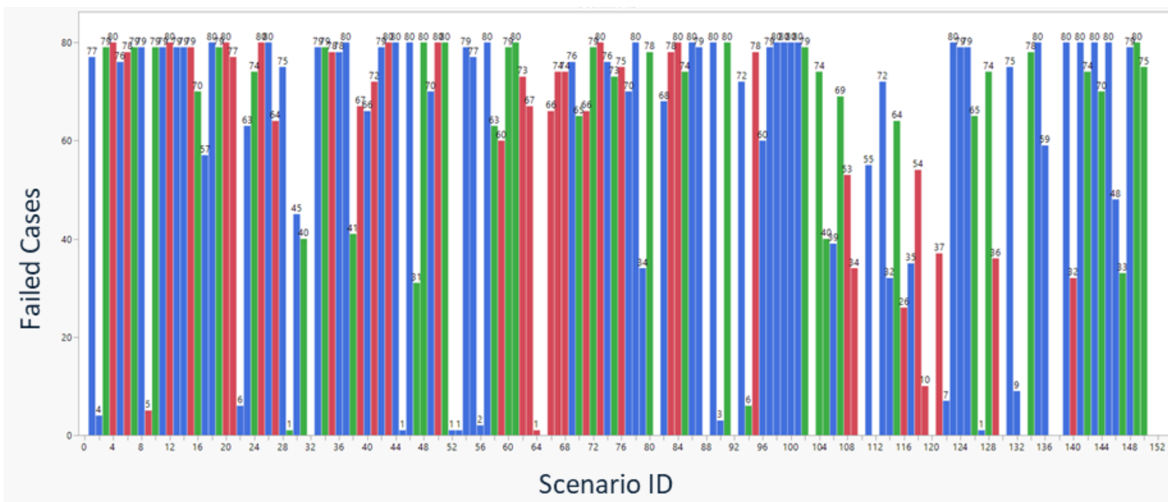


Figure 48: Failure rate by cases 80 is the maximum value, red bars are Saturation All-Out deterrence strategies, blue bars are Persuasive deterrence strategies, green bars are Saturation deterrence strategies.

Figure 49: Correlation heatmap between Aysr's and NATO’s variables.

10.2.2 Vulnerable Scenario

To find the Vulnerable Scenario, the results from JANUS were processed via the PRIM algorithm. In this step, the PRIM algorithm was used to identify the ranges of the scenario variables that create Vulnerable Scenario.

In the trade-off analysis shown in Figure 50, density is not affected excessively by coverage. In other words, the fact that the line is almost horizontal suggests that there is a good distribution of responses, and that even if many variables are left free the density remains high (>0.7). In our case, because we were interested in defining Vulnerable Scenarios, we opted to pick a point that had a high density and high number of restricted values. This allowed for finding the ranges of a wide number of variables as shown in Figure 51. As was previously done in Chapter 6, regarding the point chosen for the analysis, characteristics were reported in Table 17.

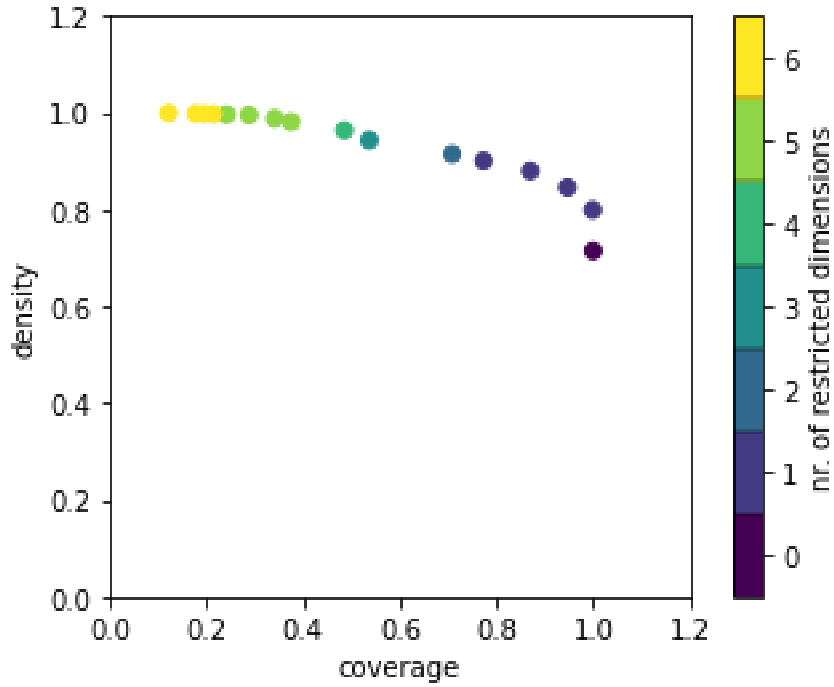


Figure 50: PRIM Trade-Off Analysis

Table 17: PRIM Iteration Results

Mean	0. 974265
Mass	0. 113333
Coverage	0. 397063
Density	0. 974265
Restricted Dimensions	6

Figure 51 is especially useful to visualize the complexity of the problem. The dots of different colors represent successful and failed scenarios for the NATO fleet. The same color code is maintained along the diagonal of the matrix where it is possible to see the distribution of the variable in successful and failed cases. Looking at the ranges of variables it is possible to see how wide they are in general, meaning that they contribute equally to Asyr's goal. The only variable that is skewed is the number of defense stations, which – without surprise – is skewed toward a high number. From the same figure it is possible to see that only two of the three strategies are considered to lead toward critical scenarios. These results are not in conflict with the correlation matrix displayed in Figure 49. The matrix showed which variable is critical, while Figure 51 is focused on identifying the range of the variables.

The information from Figure 51 are combined with the scenarios in Figure 48 to highlight the Vulnerable Scenario. These are IDs 9, 14, 26, 27, 56, 63, 77, 83, 86, 96, 111, 117, 124 and can be visualized in Figure 52.

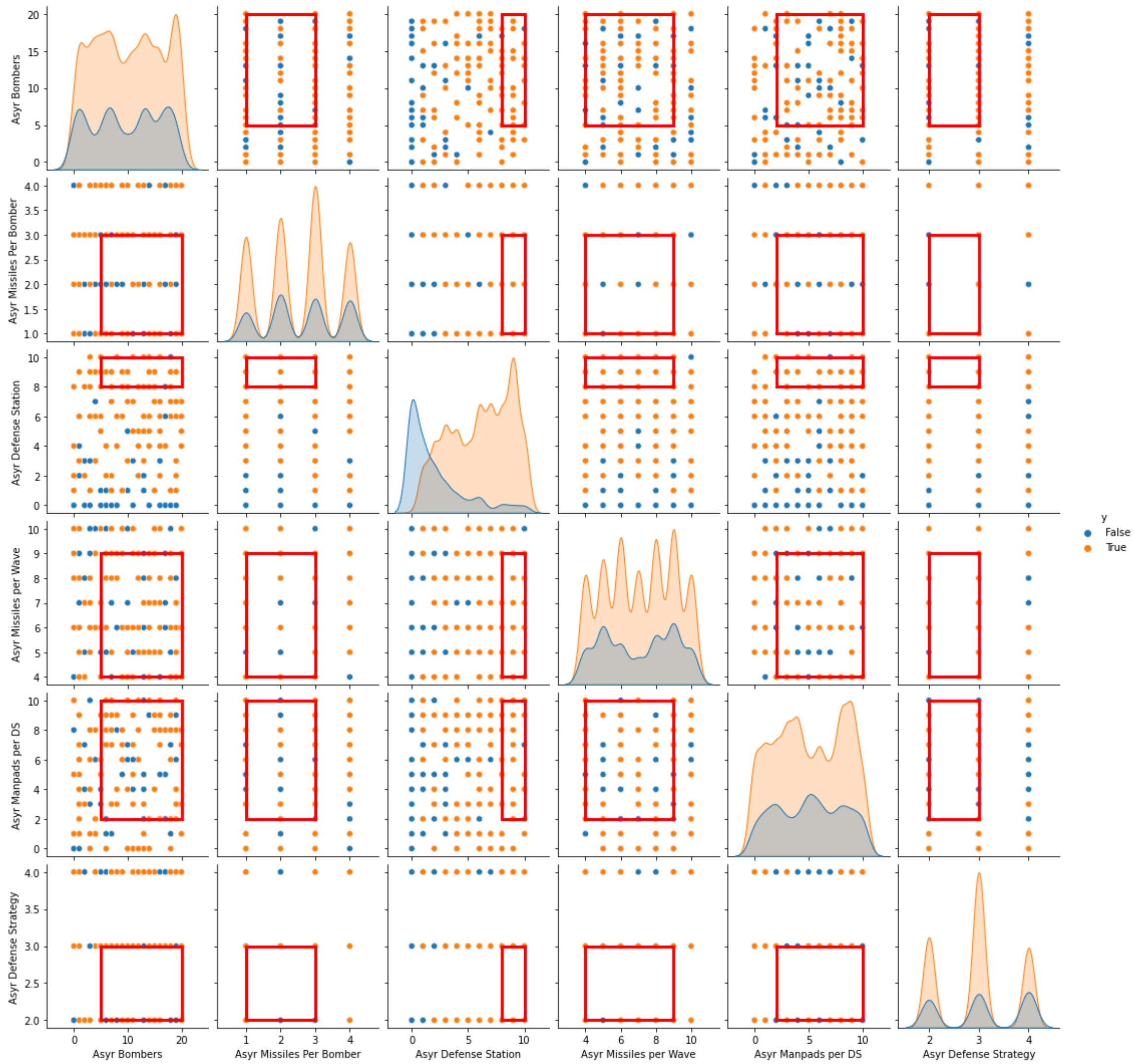


Figure 51: Results from PRIM, the red boxes represent the set of Vulnerable Scenario. In blue cases where the NATO fleet had success, in orange where it failed.

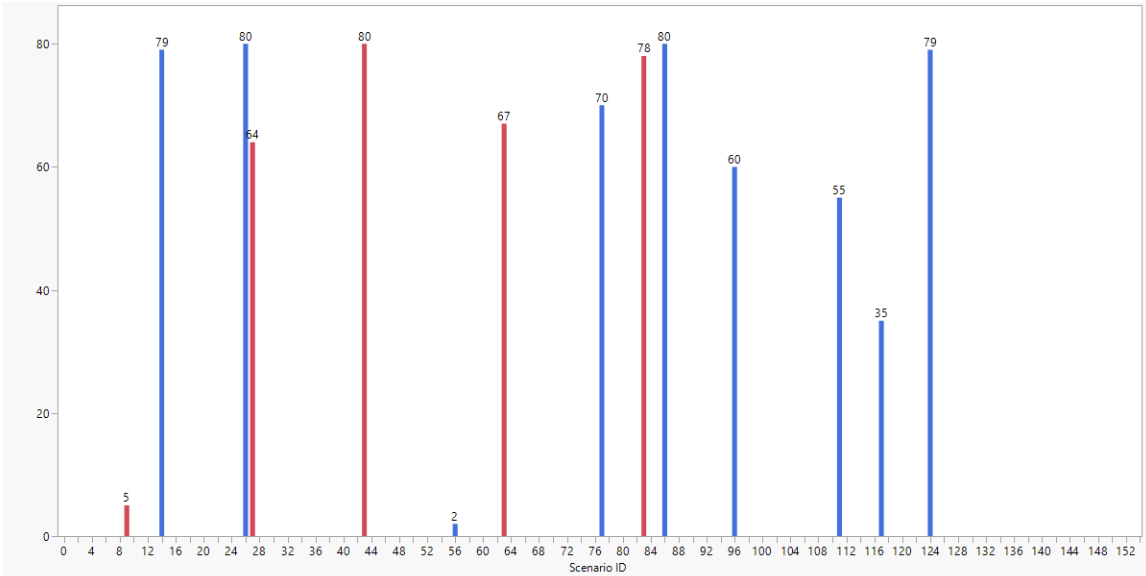


Figure 52: Vulnerable Scenario IDs identified by PRIM in point 7 (6 restricted dimensions)

10.3 Part 2: Finding Evolved Fleets

Following the methodology described in Chapter 7, the Vulnerable Scenarios discovered in Part 1 were used as a baseline to run an iterative algorithm to find evolved fleets. As was discussed before, to ensure consistency all the 80 repetitions were run again, and the random seed was fixed to the one generated in Part 1. All the information regarding the Vulnerable Scenarios and the basic fleet were transferred again from JMP to JANUS. Here, every fleet was tested and evolved according to the iterative algorithm. The iterative algorithm was designed to prioritize cheaper changes that might have an impact rather than changes requiring big ship redesigning. This was chosen following the logic that since resources are limited, it is desirable to achieve success with the bare minimum expenses. In this sense, the algorithm shown in Figure 53 was built to first update things external to the ships, like hypervelocity ammunition, short-range missiles, and naval guns. Only later to the number of iterations was capped at 30, as it was the maximum number of upgrades possible for the fleet using this algorithm.

The outputs from the simulations included not only the same outputs of Part 1, but also the newly modified inputs and the number of iterations performed. Using the inputs generated by the iterative algorithm it was possible to extrapolate how the fleet was modified. Moreover, by looking at the number of iterations it was possible to conduct a sanitation analysis on the algorithm. This analysis was performed to ensure that all the possible combinations were tried before the fleet was declared to have failed the scenarios. The sanity check proved that the algorithm performed all the required trials before outputting the fleets.

Finally, the outputs were taken and uploaded on JMP where they were checked against those of Part 1 to see how successful the evolved fleets were compared to the basic one.

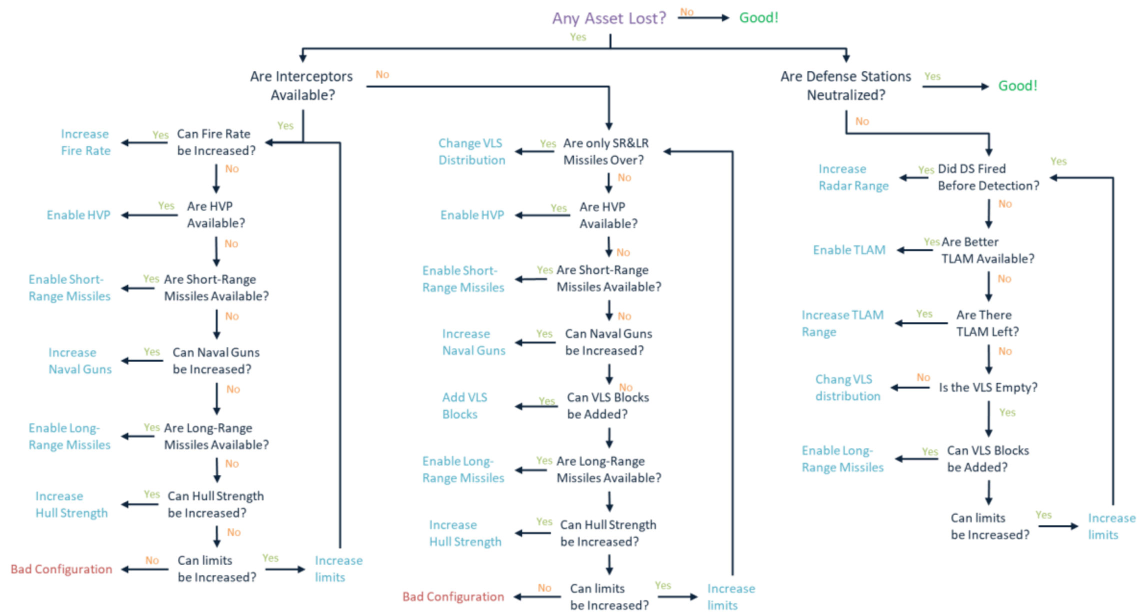


Figure 53: Structure of the iterative algorithm

10.3.1 Positive monotonic behavior verification

As discussed in Chapter 8, this check is needed to verify the behavior of the technology of interest. While it is expected that each of them can only contribute positively, this is not guaranteed. As such, in this step a verification of the monotony is conducted.

The methodology followed was the one drafted in Chapter 8:

1. 10 fleets are drafted, each with only one of the technologies of interest enabled and maximized. Differently from Chapter 8, not all the technologies were tested as it appeared clear from Part 2 that there was no need to test the VLS additional blocks, as it was never used. The whole list of fleets is reported on Table 18.

2. The created 10 fleets are included in the file with all the 150 scenarios, each repeated 80 times. To ensure consistency, the random seeds used are the ones generated in Part 1.
3. The csv file is then uploaded to JANUS, where each case is simulated and run.
4. The outputs are transferred from JANUS to JMP for the postprocessing analysis. In the results (fully reported in Appendix G), the goal is to see if some of the fleets with technology improvements have a higher failure rate than the Basic Fleet.

Table 18: Description of the technologies modified in each of the fleet used in Part 3 of the main experiment

Fleet	Modified Parameter
1	Basic Fleet as in main experiment part 1
2	Hyper-Velocity Projectiles used
3	New Short-Range missiles used
4	New Long-Range missiles used
5	Naval Guns maximized
6	Hull Strength maximized
7	Radar Range maximized
8	VLS Fire Rate maximized
9	New TLAM used
10	All technologies enabled and maximized

To understand the results obtained, each fleet was plotted against its failure rate across all repetitions in each case. In Figure 54, is reported an extract of the results of the simulation. The maximum value for each fleet is 80 – meaning that all 80 repetitions have failed. To see if the chosen technology has a positive monotonic behavior, fleets 2 to 10 are compared with the Basic Fleet (fleet 1). If the number of failed cases in one of the fleets from 2 to 10 is bigger than the number of failed cases of fleet 1, then that fleet employs a technology that has a negative effect on the fleet. Figure 54 shows a snapshot of a behavior that can be found across several scenarios. Many scenarios with a Saturation and Saturation All Out deterrence strategy – in red in Figure 54 – are negatively impacted by an increase in Naval Guns availability on the ships. Scenarios with a Persuasive deterrence strategy showed on the other hand that a higher fire rate does not help the fleet in those scenarios.

Regarding fleet 5, the increased number of failures can be justified by the fact that while the number of naval guns on the ships increased, the number of magazines was not. In scenarios where Asyr can fire multiple waves of missiles, magazines are emptied between the first and the second wave, leaving ships exposed to further attacks. Regarding fleet 8, increasing the VLS fire rate means that more missiles are shot at incoming threats. Therefore, when Asyr has the capability of firing multiple waves, NATO ships are left with empty VLS magazine and have to solely rely on their CIWS and naval guns.

Even if two variables showed a non-positive monotonic behavior, combined effect of all the other technologies. In fact, we see that fleet 10 has a 0-failure rate compared to the 76 failed cases of scenario 27 and the 50 ones of scenario 37. This was expected as the issue was addressed in Chapter 8 with the demonstration of Hypothesis 4.

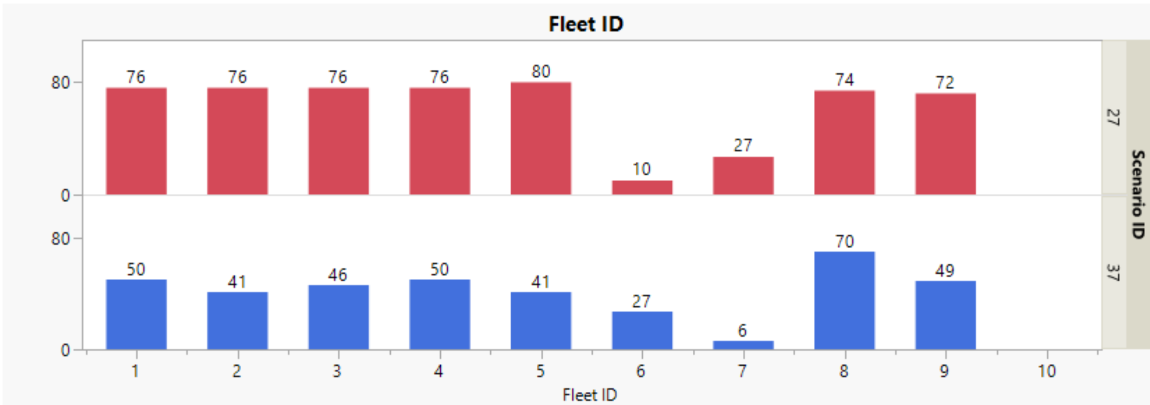


Figure 54: Detail of the results on variable monotony analysis

10.3.2 Results and discovered Evolved Fleets

The results from JANUS showed that none of the cases failed to be simulated. This is good, as it means that the model was stable and that the simulated technologies did not compromise its behaviour.

In Figure 55, quantitative way which technologies were mostly used. In this graph, on the x-axis there are all the technologies, in each column the different numbers in the internal x-axis represent the level of technology reached by the iterative algorithm. The percentages show the distribution of the different values reached across all the tested scenarios. On the vertical axis the technologies are arranged by deterrence strategy. Looking at improvements common across all strategies, it emerges a lack of interest in new TLAM and in improving the range of the radar system. A VLS blocks are not present in the figure as they were never used. Focusing now on just the Persuasive deterrence strategy it emerges an interest in new HVP ammunitions, short-range missiles, and long-range missiles. In most cases, additional naval guns are needed to support the CIWS. The CIWS is also expected to be able to fire at higher fire rates (as this variable represents the time between firing, the lower the better). Upgrades to the hull plating are generally needed to increase the survivability of the ships.

Regarding the Saturation strategy, fewer modifications are needed. In most cases, there is no need for new short-range and long-range missiles. Similarly, new naval guns seem to be obsolete as they show improvements only in 16% of the cases. The ship structure does not require additional strength in over 85% of the cases. The only two areas in which we see some needed upgrades are HVP ammunitions and fire rate which are upgraded 30% and 34% of the cases, respectively.

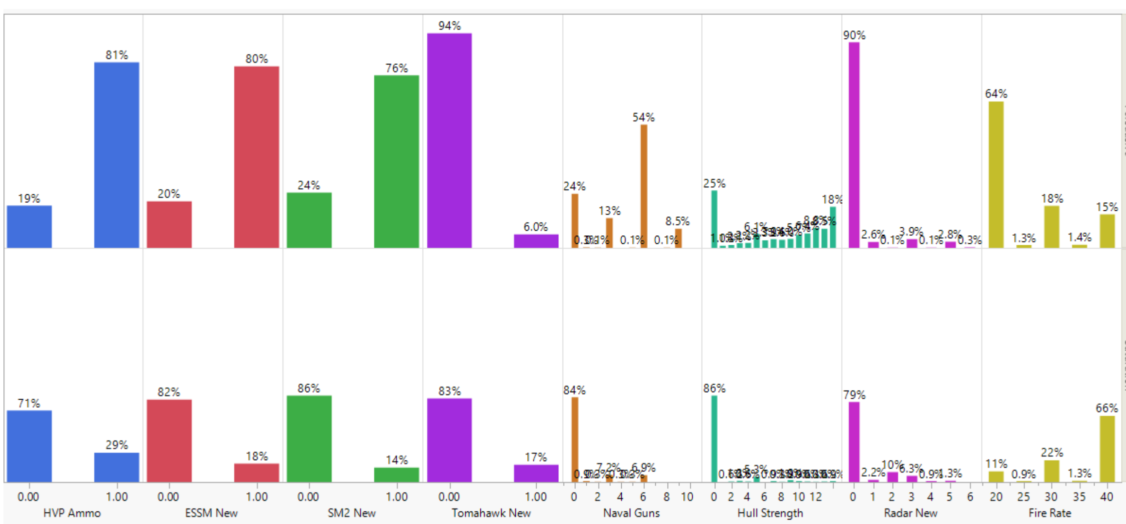


Figure 55: Representation of which technologies were upgraded, top row shows persuasive deterrence strategy, bottom row shows the saturation deterrence strategy. On the y-axis, percentage of usage.

The pools of technologies described above can be aggregated using the methods demonstrated in Chapter 8 to derive the evolved fleets which will be used in Part 3. These fleets are reported in Table 19, duplicated fleets like the one generated in IDs 77, 96, 111 will be removed to reduce computational efforts.

Table 19: List of fleets generated after as result of the Part 2 step of the methodology

Scenario ID	*Max of HVP Ammo*	*Max of ESSM New*	*Max of Long-Range Missiles*	*Max of Naval Guns*	*Max of Hull Strength*	*Max of Radar New*	*Max of VLS Distribution*	*Min of Fire Rate*	*Max of Tomahawk New*
14	1	1	1	6	14	5	1	20	1
26	1	1	1	6	14	5	3	20	1
27	1	1	1	6	10	5	3	20	1
37	1	1	1	9	14	5	2	20	1
43	1	1	1	3	12	7	1	30	0
63	1	1	1	6	14	0	1	20	1
77	1	1	1	6	14	5	1	20	1
83	1	1	1	6	10	2	3	20	1
86	1	1	1	6	14	6	3	20	1
96	1	1	1	9	14	5	2	20	1
111	1	1	1	9	14	5	2	20	1
117	1	1	1	6	14	3	2	20	1
124	1	1	1	6	14	3	1	20	1

From the data in the table, it emerges how the cheapest modifications while more expensive ones such as Hull Strength see more variability. Regarding Hull Strength, values is the maximum of the range. In fact, as few cases have been successful only at 20+ iterations, through all the technologies, leading the fleets to have very high values of Hull Strength and Naval Guns.

From Figure 56, it is possible to see how the success rate was much higher compared to the 32.6% achieved in Part 1. The secondary objective to destroy all defense stations was also achieved most of the time. As this was a secondary objective, the algorithm tried to ameliorate the fleet in that area only when the fleet was failing the primary objective. This means that fleets in those scenarios going through the iterative algorithm only a few times might not have had the occasion to strengthen offensive capabilities enough. Finally, the blue area are the cases in which the fleet still failed to survive. In all those cases, after 30 iterations, only one frigate is destroyed. Comparing this result to what was discovered in Part 1, there is a general reduction in losses. In fact, even in failed cases where previously 6 or 7 ships would have been sunk, now that number is reduced to 1, proving that even if the main objective is not fully achieved, the iterative algorithm provides a strong contribution toward the fleet's resilience. This should not be seen as a change in success criteria, but as a fact to be stated, i.e., failures are less heavy in terms of losses. In this sense, this result is a byproduct of the methodology, showing that if the decision maker is interested, partial success is possible, and that the methodology is able to capture it.

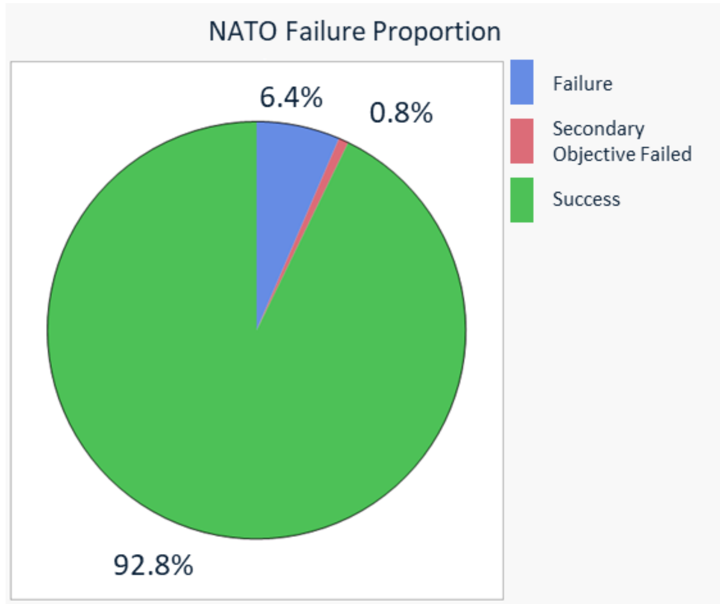


Figure 56: Success case distribution in relation to objectives.

Finally, in Figure 57 there are those scenarios which were not able to be fixed by the Evolved Fleets. As reported before, even if those were failed scenarios, the degree by which they failed was reduced. By comparing Figure 57 with Figure 52, it is possible to see how the failure rate changed in each of Vulnerable Scenario thanks to the use of the technologies found through the iterative algorithm. The positive effect of the found technologies is particularly evident in all those cases which went from a failure rate of 80, over 80 simulations, to less than 5 failed cases. This demonstrates how new technologies make the difference in mission's success and how, without these improvements, the fleet will not be able to overcome challenges imposed by adversaries in Vulnerable Scenarios.

It is interesting to note that almost all the Saturation defense scenarios were resolved, while Persuasive strategy ones remained the most critical. It is worth mentioning that following the result of this part of the experiment an analysis of why this deterrence strategy proved to be more challenging for the NATO fleet was conducted. The analysis showed that those scenarios in which the fleet failed have a concentration of defense stations on one of the two sides of the channel. This caused one side of the fleet to take most of the hits, saturating de facto the defenses of one or two ships. The other flank of the fleet on the other hand was not really exposed to challenging threats and it was too far to support the defense of the ships under attack.

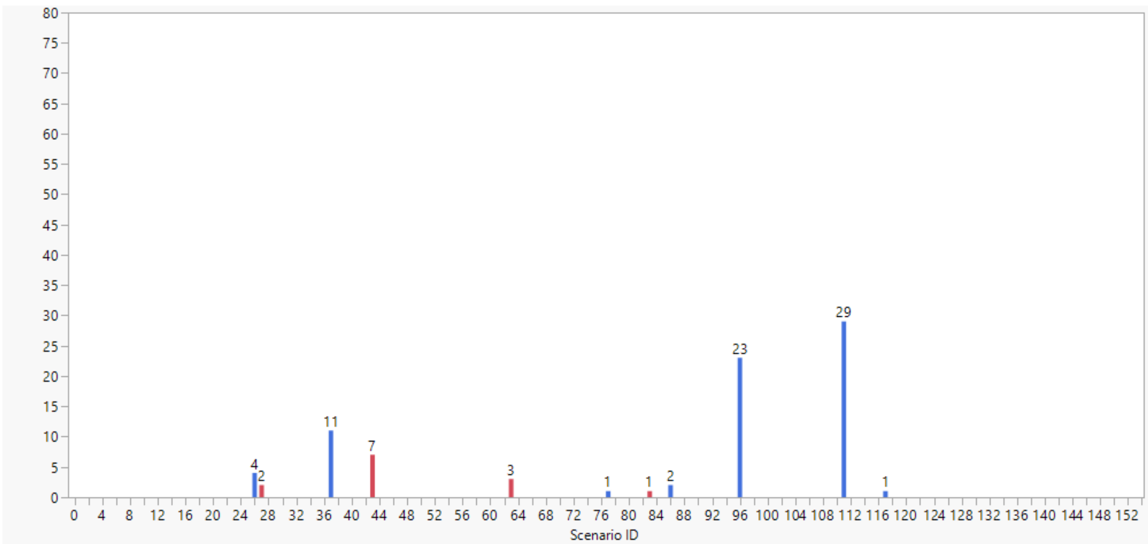


Figure 57: Scenarios in which the fleet still fails after the iterative algorithm (in red saturation deterrence strategies, in blue persuasive deterrence strategies).

10.4 Part 3: Finding Robust Fleets

Part 3 of the experiment focuses on two key issues. First, it looks at identifying which variables have a monotonic positive behavior, and the related consequences. Secondly, it focuses on identifying the robust fleets among those generated in Part 2, providing a conclusion to the experiment.

10.4.1 Finding robust fleets

The last step of the experiment is to identify and then rank which fleet was discovered in Part 2. Of those originally 13 fleets only 10 were kept as unique. The value level of each technology is reported in Table 20. As the simulation followed the same process as that in the previous paragraph, the 10 fleets were taken and tested in all 150 scenarios – each repeated 80 times. The same random seeds were kept ensuring consistency and repeatability of the results. After the csv file with the 120.000 cases was transferred to JANUS, the agent-based modeler took care of performing the simulation. The results were provided again as another csv file, this was later uploaded on JMP where results were analyzed. Table 20 also shows the failure percentage of each fleet. From the results, it emerged that Fleet 4 was successful on the 99.9% of the cases, with only 17 failed cases out of the 12.000 in which it was tested. Fleet 1 followed with 99.7%, then Fleet 10 with 99.1%, Fleet 9 with 98.9% to conclude with Fleet 6 and its 96.9% of success rate. The other 5 fleets had a success rate below 80% and as such they were later discarded.

Table 20: Fleets tested and their technologies

Fleet ID	*HVP Ammo*	*ESSM New*	*Long-Range Missiles*	*Naval Guns*	*Hull Strength*	*Radar New*	*VLS Distribution*	*Fire Rate*	*Tomahawk New*	*Success Rate*
1	1	1	1	6	14	5	1	20	1	99.7%
2	1	1	1	6	14	5	3	20	1	77.6%
3	1	1	1	6	10	5	3	20	1	77.4%
4	1	1	1	9	14	5	2	20	1	99.9%
5	1	1	1	3	12	7	1	30	0	77.5%
6	1	1	1	6	14	0	1	20	1	96.9%
7	1	1	1	6	10	2	3	20	1	77.3%
8	1	1	1	6	14	6	3	20	1	77.4%
9	1	1	1	6	14	3	2	20	1	98.9%
10	1	1	1	6	14	3	1	20	1	99.1%

Figure 58 confirms that the Persuasive deterrence strategy has a higher failure rate for the NATO fleet. In fact, we see that most failed cases are within the Persuasive deterrence strategy. From the Vulnerable Scenario discovered in Part 1, this had to be expected, and the reasons why this happens, and how this is linked to the simulation were reported at the end or Part 1.

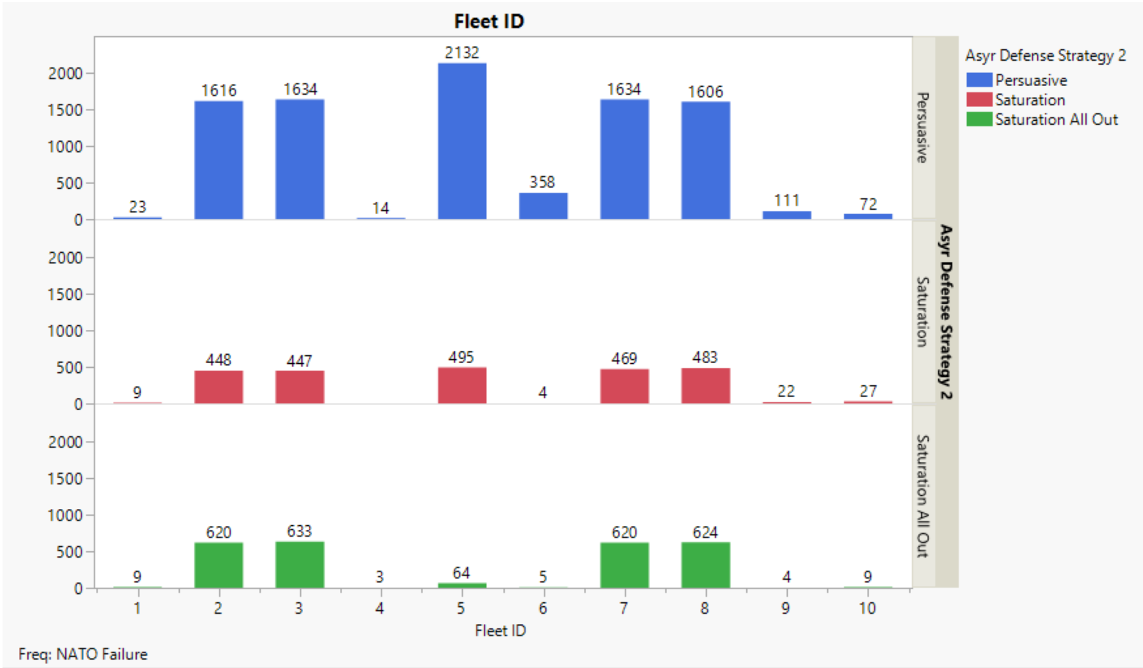


Figure 58: number of successes per fleet per Asyr deterrence strategy. The maximum value on the y-axis is 4000, as the maximum number of cases per Asyr's strategy.

Looking more in depth in the cases that were failed by these fleets as reported in Figure 59, they are spread across multiple scenarios, with predominance of cases 30, 33, 55, 60, 93, 96, 102, and 139. One common factor of the 8 scenarios mentioned is the high number of defense stations (above 8), which is consistent with what was found in the Vulnerable Scenarios in Part 1 of this experiment. Another commonality is the high number of MANPADS (above 5) around each defense station. These two facts combined show how on one hand there are many targets for the incoming NATO fleet, but on the other, those targets are also very resistant. This means that the NATO fleet is in the hard position of having to distribute its fire power across many resisting targets which at the same time have a large fire power themselves.

It is also interesting to note that the number of missiles, and the number of waves that each defense station fired, were not common factors. This shows that the distribution of incoming threats, and therefore their ability to saturate parts of the fleet's defenses, proves to be a tougher challenge for the fleet rather than the number of total missiles distributed across a longer period. A logical explanation of this is that the layered defenses employ both missiles and CIWS. Missiles independently target each threat and, depending on the fire rate, multiple missiles can be shot at the same incoming threat. On the contrary CIWS can only fire one incoming threat at a time, usually at a much shorter range. Targeting two threats in two completely different directions is not possible for the same CIWS. While it is true that ships have multiple CIWS covering all the possible incoming directions around the ship, it is also true that CIWS can get saturated when too many threats are incoming from the same direction. For all these reasons, these results are consistent with what was seen in the past with ASCM [97], and with what we will most like see in the future with aerial swarms of autonomous vehicles [98].

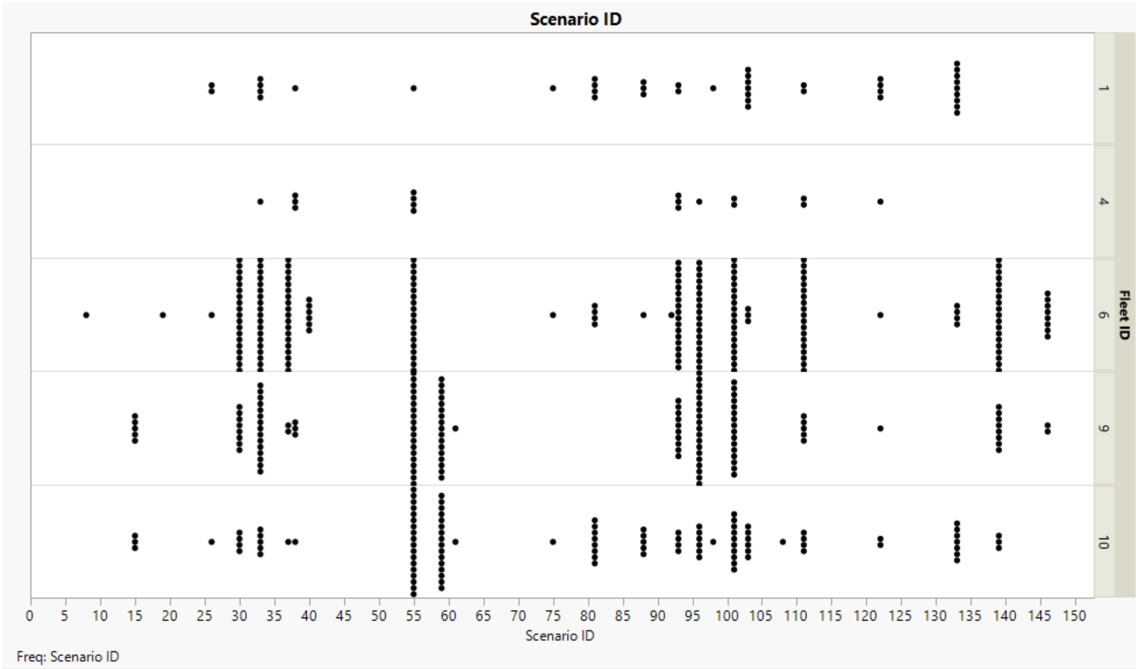


Figure 59: Distribution of failed cases across the 5 fleets with a success rate higher than 80%.

10.4.2 Conclusions

To conclude this chapter, and the demonstration of the full methodology, the only thing left to do is to rank the fleets highlighted as robust in the previous paragraph. To further reduced the 10 fleets studied, an 80% success rate as acceptability criterion was imposed on the fleets. In the case of this experiment, the selected value separates the pool of fleets into 2 equal divisions

Table 21 reports the summary of the main characteristics of the 5 fleets and their relative rankings. Two main characteristics were chosen to rank the fleets: their success rate and the number of upgrades performed to each fleet. The reason the success rate was chosen is straightforward: we are interested in the most robust fleet across all scenarios and repetitions without considering economic aspects. In that case, Fleet 4 would be the one to be picked as it has the highest success rate. On the other hand, if resources are constrained it might be of interest to look at the number of upgrades (i.e., investments) needed per fleet. This metric is in tension with success rate; therefore, it is not a surprise that the first choice in this case would be Fleet 6. By screening the original 10 fleets imposing a minimum success rate it was ensured that even the fleets chosen because of their lower modification number have a satisfactory success rate.

Table 21: Fleets ranking and main characteristics

Fleet Id	Success Rate	Success Ranking	Number Of Modifications	Modification Ranking
1	99.7	2	32	4
4	99.9	1	35	5
6	97.1	5	27	1
9	98.9	4	30	3
10	99.4	3	30	2

CHAPTER 11. Conclusions

The proper function of man is to live, not to exist. I shall not waste my days in trying to prolong them. I shall use my time.

Jack London

This thesis started underlining a problem that today is emerging increasingly: the shifting of the security environment toward a more volatile, uncertain, complex, and ambiguous one. It was highlighted how to compensate for this shift there could be solutions employing different means, ways and ends. One of the first assumptions was that through this research, a solution could be identified by looking at means, specifically, there is interest in looking at means within the naval domain. It was identified, differently from all differently from all the other domains, that the sea domain offers an ample array of relevant use cases and scenarios that will continue to be pivotal in shaping the future of the security environment. Therefore, assessing the fundamental roles of these scenarios will support the field. In the end, through the development of the methodology showed in the Figure 4 the main motivational questions were answered:

- Motivation Question 1: What ship taxonomy can be used to better understand high level interactions among subsystems?
- Motivation Question 2: What are current investment decision making methodologies used by the DoD?
- Motivation Question 3: How are assets currently quantitively compared in investment methodologies?

Throughout the course of this thesis a series of research questions were formulated and answered in different chapters. Table 22 shows a summary of these questions and the locations where they were addressed.

Table 22: Summary of Research Questions and locations

Research Question ID	*Research Question*	*Chapter*
Research Question 1	Can new taxonomies of ships help in increasing the understanding of high-level interactions among components?	2
Research Question 2	What modelling techniques can be used to quantitively simulate naval future scenarios?	3
Research Question 3	What technique should be used to find vulnerable scenarios in a large dataset with deep uncertainty on the future evolution?	3
Research Question 4	What Scenario Discovery method can be used to find Vulnerable Scenarios for naval fleets in a credible and rapid way?	3
Research Question 5	What tool can be used to select sets of naval technologies to enhance fleets' long-term robustness?	4
Research Question 6	Which hybrid approach can be used for quantitatively selecting technologies to invest in a naval fleet in a credible, practical, and rapid way?	4
Research Question 7	What happens if technologies effects are not positive monotone?	4
Research Question 8	Which criteria should be used to select a reduced number of fleets to be further evaluated for robustness?	5

A variety of gaps emerge when addressing the means. The first one is the way new assets are planned. Different bottom-up and top-down methods are present but the reconciliation between these is not always straightforward, therefore it can happen that assets that satisfy planners' needs end up not satisfying operational requirements. Part of the efforts to cover this gap is reflected in the way used how to decompose assets: taxonomies.

In Chapter 2 the goal was to answer Motivation Question 1, to do so multiple taxonomies models were described, highlighting why each of them taken individually would not satisfy the need to comprehensively describe complex systems as ships. The conclusion of the chapter demonstrated, by answering Research Question 1, how hybrid taxonomies can solve the problem and can allow clear and quantitative comparisons among assets.

Being able to compare assets is another one of the issues that has been tackled in this thesis as described in Motivation Question 3. In fact, to provide quantitative solutions to show which means can support a stabilization of the security environment, it is necessary to be able to quantitatively compare and model those means. The comparison is done by using hybrid taxonomies, but regarding the modelling part there are different options. Among all the options presented to answer Research Question 2, it was chosen to use agent-based modelling due to its versatility and its ability to well describe many independent assets collaborating for one outcome.

All these pieces are important to set the stage for the two key deliverables of this methodology: the future scenarios that will be challenging for the assets we are interested in investing on, and the set of technologies that will help those assets maturing to the point of overcoming those found scenarios. To find future scenarios the investigation started by formulating Research Question 3 and by selecting Scenario Discovery (SD) as a technique to answer that question. SD can use different statistical tools, Research Question 4 was formulated to identify which one was better suited to find Vulnerable Scenarios in a naval context. The choice fell on Patient Rule Induction Method (PRIM) as tool to analyze scenarios and frame relevant scenario boxes in the design space through a study on the ranges of the used variables. In performing the literature review in Chapter 1 through 5, it emerged that PRIM was never used in conjunction with agent-based modelling to find naval Vulnerable Scenarios, therefore a window of opportunity to contribute to the field was identified.

The second key deliverable is the sets of technologies allowing fleets to overcome the previously discovered Vulnerable Scenarios. In this effort, hybrid taxonomies and the use of a DoE proved to be essential as different technologies contributed to fleets at distinct levels. To maintain a broad vision of the problem, and to be able to test multiple technologies (>10), a hybrid approach that answered Research Question 5 was developed. The requirements for the hybrid approach were investigated in Research Question 6, this had to reflect the need to start with a fixed starting point but then to allow the fleet of interest to be upgraded through a selected group of technologies. For the static part, it was chosen to use a space filling DoE, more specifically a Latin Hypercube Design, while for the dynamic part a new method had to be developed. To be consistent with the use of SD in Chapter 3 several methods used in the past with SD were considered. None of the methods or existing approaches satisfied the requirements needed to support the objective of this thesis, it was decided to develop a new method. This involved using a signpost and trigger system that was able to track the evolution of the variables of interest from one simulation to the next, then update the inputs of the subsequent simulation. The selection of the new inputs was done by tying the signpost and trigger system with a decision-making tree that provided for the new variable's values depending on the output of the previous simulation. In identifying the different set of technologies, it emerged the possibilities of those technologies not contributing only positively toward solving Vulnerable Scenarios but also opening new ones. Therefore, first Research Question 7 was posed, then, as part of the method to solve this issue, the experiment in Chapter 8 was run to demonstrate a necessary checkpoint determining the positive monotony of the variables.

With all the information needed placed together it was possible at this point to create a new methodology that was able to satisfy the original research objective of developing a procedure to support concurrent trades-offs among naval assets and technologies, to assist investments on new long-term maritime technologies.

The newly created methodology was divided in three parts: the first one with the goal of identifying Vulnerable Scenarios using PRIM, the second one tasked to find the Evolved Fleets using the hybrid approach created ad hoc, and the third one to find the Robust Fleets by aggregating the results obtained from the previous parts. Each part was dedicated to an experiment so that hypotheses and assumptions could be validated before testing the whole methodology as one block. In Chapter 6 an anti-submarine warfare (ASW) use case was set up to identify Vulnerable Scenarios for a single ship and its helicopter. This use case was built as the foundation for the experiment performed in Chapter 7. The output of this was the identified scenarios of different technology strategies that enabled the fleet to succeed. Chapter 8 was dedicated to verifying the positive monotony of the variables, to find if possible new Vulnerable Scenarios opened, and to see if among the technologies tested some can compensate for the eventual non-positive monotone technologies present. Finally, the experiments in Chapter 9 demonstrated hypothesis 6 and answered the last research question, Research Question 8, testing how the fleets could be aggregated and reduced to find the most robust combinations within minimal computational time. In this chapter the upgraded fleets from Chapter 7 were compacted so that from each Vulnerable Scenario only one fleet was output. At the end of chapter, the selected fleets were compared and tested across all the scenarios to find those with the highest success rate.

The final demonstration that supported the research in this thesis tested the whole method as a single block in Chapter 10. In this chapter a different, and more complex, use case was selected: a fleet of 10 ships had to approach a group of islands to perform a Non-combatant Evacuation Operation, in approaching the islands the fleet had to suppress several fixed defense stations, some of which not known, and neutralize an aerial attack. The methodology was executed in its entirety and the robust fleets were promptly identified. The experiment was a successful demonstration that the whole methodology can work with different use cases and that it satisfies the research objective.

In conclusion, this thesis has been developed in the years with idea in mind that it could support the work done by planners in the S&T community. The results obtained as deliverables (i.e., Vulnerable Scenarios and Robust Fleets) are not finite products for someone who is looking at investments, but they could support planners in evaluating different, unexpected, and yet quantitative scenarios. Although these are partial results, they can be used to reduce the lead time in identifying which investment will work better in an uncertain future. Moreover, the outcomes from the partial experiments and full methodology demonstrate the ability to focus high-detailed modelling efforts to certain portions of the design space, while allowing decision makers to have abroad outlook of what the future holds. With all of this in mind, entities that could benefit from this work include bodies like the Joint Requirements Oversight Council (JROC), to support its role in validating joint warfighting requirements, and Acquisition Offices focused on identifying future investments.

In the future it would be interesting to integrate the methodology developed in this thesis with already existing ones like JCIDS. This could allow further streamlining of the modelling efforts and could potentially increase the number of scenarios tested. Moreover, it would be interesting to integrate some higher fidelity modelling abilities to study specific effects that might be of interests for decision makers. This could enable further tailoring of the methodology to help decision makers understanding more in depth how technologies of interest work, while keeping in mind the large investment effort. Looking at how to ameliorate the modelling structure, it could be possible to integrate better degrading models for ships. This would allow to shift away from the binary condition used in which ships were either perfectly functioning or destroyed. Further modification to the model could involve differentiation of chaffs and flares in the countermeasure array available to helicopters. By doing so, it could be possible to investigate more type of missiles. Finally, future works could include expanding above the naval domain, to include segments from other Armed Forces. This would show not only the potential of joint operations, but also the benefits of streamlining technology investments across all branches of the Armed Forces.

Although this work was focused on a military naval perspective, the methodology can be broadly applied to many different domains. With this work the author hopes to have satisfied the curiosity of those seeking different options for finding unexpected scenarios that might affect their fleets, and to contribute to the future of scenario problem solving in the naval domain.

Appendix A. List of Maritime technologies of future interest

A.1 Surface Vessels

Throughout the next few paragraphs, the goal is to show several ships' subsystems in which currently there is interest on investing in technologies to counter A2/AD bubbles, and possibly more. The division in subsystems follows a functional decomposition.

Surface ships can be of different types, from corvettes to aircraft carriers and from Off-shore patrol vessels to amphibious assault ships. In the following subsystems decomposition, all the subsystems were abstracted removing them from their original ship context. This does not mean that any ship can mount any technology, but that at an abstract level different ship classes can be compared on their basic elements.

i. Warfare Subsystem

One of the main problems of entering in a contested A2/AD area is surviving the incoming attack. Differently from the past when the attacker had the advantage (e.g. Falkland war), the defenses here will have the advantage of larger magazines and more fire points. Moreover, the attacker after surviving the first wave of incoming strikes must be able to have enough magazine to fire its own attack versus the defense stations. This means that Vertical Launcher Systems (VLS), that cannot yet be reloaded at sea, must be optimized to provide ship safety while also providing an effective offensive measure. Currently on US

CG and DDG-51 only 25% of the VLS magazine is used for TLAMs, this number increases to 40% on the 3 DDG-1000 the US currently has [99].

To increase the survivability of the fleet studies have demonstrated the need of new weapons that, together with an upgraded battle management system, can provide greater resiliency and a lower dependency from magazines. This is achieved by complementing hard kill systems like missiles with soft kill ones like High Power Microwaves (HPM) weapons, Direct Energy (DE) laser systems, and jammers for electronic warfare.

HPM weapons are a promising technology combining the capability of hitting multiple targets at the same time, in all weather, with an endless magazine. HPM weapons work by sending a narrow beam of high power microwaves that interferes or cause damage by inducing currents in the targeted circuits leading to overheat failings similar to the blowing of a fuse [100]. Currently DoD is investing on a prototype called CHAMP. DE systems on the other hand are affected by the humidity level of the atmosphere making it not optimal for operations in the maritime environment. Both systems though require a certain time locked on target to achieve their effect, this disadvantage is countered by the fact that all prototypes so far have proven that multiple beams can be fired from a single system. Another negative aspect is that certain weapons, like ASCM, are hardened to resist the high temperatures of supersonic travel, therefore they will likely be not affected much by HPM and DE weapons. Jammers work in a different way as they try to “blind” the incoming missiles by disrupting and confounding its navigation systems ad seekers. Jammers are also used as decoys as they can simulate the IR and radar signature of ships of different sizes luring the incoming missile away from the HVU.

Hard kill systems include naval guns, missiles, and in the future also railguns. Naval guns are now seeing investments in better munitions, specifically on Hyper Velocity Projectiles (HVP) which can reach speed of Mach 3 when fired by powder guns. HVP have small control surfaces to correct their course more accurately, compensating for small errors in target location. Electromagnetic Railgun (EMRG) can use HVP as well providing more range, up to 40 nm, but they require more power that currently only DDG-1000 are able to provide. Moreover, the fire rate is slower than other naval gun systems, making them not as appealing [101]. Finally, to increase the VLS magazine a possible solution would be to start "quadpacking" missiles like it's currently done with ESSMs, this will reduce the range of some of the offensive missiles moving them away from being standoff weapons. Looking at a different concept, some have suggested investments on magazine ships, which should be focused on providing only a large number of VLS cells, between 256 and 512, while relying on other ships in the fleet for defense [23].

Torpedoes are of course part of the warfare subsystems, but they are not as critical as other weapons in entering or in surviving an A2/AD bubble.

ii. Sensor & Processing Subsystem

An important industrial effort in the past years has been focused on increasing detection capabilities of ships. This has been done by improving sensors on board and by augmenting radar's range resolution. Upgrades are also achieved via software as done in the case of the DDG-1000. Cpt. James Syring, DDG-1000 program manager, after cancelling the order for the S-band SPY-4 volume search radar stated the following: "We don't need the S-band radar to meet our requirements [for the DDG1000]" and "You can meet [the DDG-1000's operational] requirements with [the] X-band [radar] with software modifications” [102]. On the other hand, other S-band 3D radars are currently being developed, among them the next generation of US ships will be using the AN/SPY-6. This is an air and missile defense electronically scanned array 3D radar which will also provide for periscope detection. Table 23 provides a list of radars on an Arleigh Burke destroyers (DDG-51) and their main functions. This list is provided to show how different technology investments could be performed in each of those categories to increase DDG-51 capabilities.

Table 23 Types of radars on DDG-51

Radar Name	Function
AN/SPY-6(V)	Air and Missile Defense Radars used mainly as air search radar. It replaces the AN/SPY-1D in all Flight III ships.
AN/SPG-62	Illumination radar for semi-active radar homing missiles.
AN/SPS-67	Dedicated surface search radar for finding and tracking ships and sea skimming objects.
AN/SPQ-9B	Replaces the AN/SPS-67 in some ships, it combines surface search and fire control capabilities. It is optimized against stealthy fast approaching assets like cruise missiles.

Sperry Marine BridgeMaster E	Used as search and navigation radar, it replaces the AN/SPS-73 as it has greater at-sea reliability and can detect sea skimmers moving at speeds up to 600 knots.

Sonars are also a critical part of the ship's sensor suite. Sonars can be both active and passive, can be mounted on the hull – usually at the bow to reduce noise – or towed about a mile behind the ship. Continuing the example of the Arleigh Burke destroyer, which is not specialized in ASW, there are 2 sonars: the AN/SQS-53C active/passive sonar mounted in the bulbous section of the bow, and the AN/SQR-19B1 passive towed sonar. The latter is now being replaced by the new TB-37U which adds the advantage of being able to operate both passively (just listening) or actively (emitting sonar pings and listening).

All these systems are controlled by the battle management computer, the brain behind the full warfare system that connects sensors to weapons through designed tactics and responses. The Aegis Combat System can be defined as the pulsing heart of the defense system as it connects the information coming from all the different sensors in a coherent battlespace to show the status of the ship and the battlefield in general. This system has been successful because of its open architecture that allows different components to be integrated, allowing ships to share and gather data with other Aegis ships. While the Aegis system is focused mainly on air and missile defense, subsurface combat systems have now started to be integrated. Aegis has the capacity to match automatically the most appropriate response to the incoming weapon, this is done following doctrines which are hard coded in the system. Usually rules involve understanding the threat, prioritizing threats if multiple are detected and then assigning the response to different defense systems. Future fire control systems should increase their agility in responses and automation as incoming threats are expected to be coming at much faster speeds and with harder detection patterns like coordinated sea-skimming attacks. Moreover, many are claiming that ships' defense bubbles should be reduced in size to allow more short and medium range interceptors [99] – like ESSM, which are cheaper and available in greater numbers than SM-3 –, HVP from naval guns, and soft kills systems. All these systems can usually engage at around 40nm, offering fewer encounters with the incoming missiles than the conventional approach of neutralizing incoming threats as far as possible from the ship. This is a shift in culture which should be supported by technology investments on the fire control system.

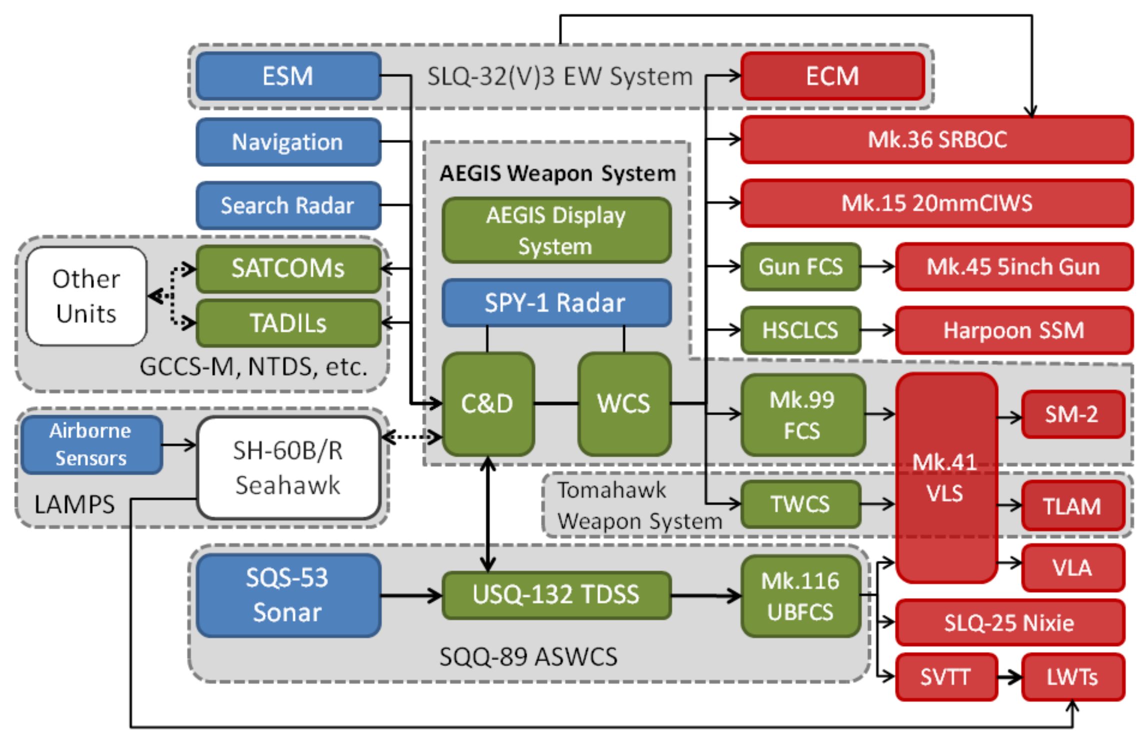


Figure 60: Schematics of the Aegis Combat System; sensors are in blue, controllers in green, weapons in red and external units in white.

iii. Communication Subsystem

The communication layer is important to maintain contact with command centers, with other ships and aerial assets in the fleet. In general, assets with an air wing will have a tactical air navigation system (TACAN) for assisting helicopters in take-off and landing operations and to inform them regarding their distance and bearing from the ship. Important communication systems include SHF satellite terminal, various links (4, 11 and 16) for tactical communications and link 22 for enabling higher data transfer. To reduce the risk of being intercepted, the Office of Naval Research (ONR) has started working in what they called Tactical Line-Of-Sight Operational Network (TALON) [103]. TALON is a laser communication system that allows ships to communicate at gigabyte speed without the risk of being intercepted. In fact, laser beams, differently from radio frequencies, are very narrow minimizing the risk of being intercepted as sender and receiver must be in line of sight to achieve communication. Further application of this technology could be considered for allowing satellite-ship communications, this would be a breakthrough for secure communication in contested areas, as A2/AD bubbles. In this field research is still going on as the maritime environment, given the high relative humidity, the constant rolling and the weather variability, poses a challenge to this type of communications [104].

iv. Structure

One of the technology trends that has been growing in the past years is the concept of modularity and flexibility. Modularity can be achieved in 3 different ways [105]: by having common modules that can be used on multiple ship classes, by using self-contained modules providing plug-and-play capabilities, or by employing modular installations on the sea frame that can allow different module packages with different functions to be installed on a standard interface. An example of the first approach is the design of a common medical facility that has standard dimensions and standard interfaces; by doing so in the design cycle the ship needs only to have a dedicated space facility and no further design would be needed. Self-contained modules are useful as they just need to be plugged in to provide the desired capability. An example of this approach is the VLS. A VLS can be changed and prepacked with any missile configuration without changing the form of the launcher. VLS can be used to accommodate not just missiles but also UAVs and torpedoes.

The third concept of modularity is the installation of modular payloads on the ship. In this case there is a full separation between the ship and the payload. The ship is required to have well defined standard interfaces where different payload packages can be installed to perform certain missions: ASW, MCM or ASuW. The US Navy tried to implement this concept on the LCS program, and while in the end it did not work well for the USN [106], other navies are still studying and implementing it. Modularity in fact offers the opportunity to repurpose ships when needed, tailoring them to specific expected threats. The USN is still partially looking at this in the FFG(X) program where in the RFP it was made clear that the ships must be upgradable without the need of dry docking or hull cuts [107]. Of course, this is not the quick module swap that was originally envisioned for the LCS, but this hybrid approach could result beneficial given the uncertainty of the current *security scenario*. It is possible to summarize modularity as the capability of a ship to adapt new technologies and perform different missions at high standards.

Flexibility, on the other hand, is the capability of changing not only the module but also the interface at ship level [105]. One way of doing it is by having flexible infrastructure that evolves with the needs of the ship, as done on Ford class. On the CVN-78 in fact, the internal bulkheads are mounted on track allowing different sizes and configurations of rooms while standard connections are provided for power, cooling and computers. Alternatively, ships can be built slightly larger than needed allowing more space for further upgrades, as done on the Ticonderoga-class cruiser.

v. Propulsion Subsystem

In 2016 MITRE conducted an analysis on the future fleet composition [23]. Among the different suggestions to counter A2/AD they showed the need of increasing the number of air wings to reduce the attrition of incoming air forces and missiles. One way to do this is to increase the affordability of carriers by developing a scaled and cheaper version which can support STOVL aircrafts (e.g F35-B). This new carrier should be either based on the America class LHA or a completely new design, but in both cases, it should use a conventional diesel propulsion system. Investing in a different propulsion system will allow to reduce costs from the original nuclear propulsion design. Affordability is a key aspect in future scenarios where more assets will be need project more forces.

vi. Air wing, Boats & Landing crafts

It was already mentioned how air tankers will be critical to keep the carriers away from coastal defenses while fighters are projected in. For this reason, one of the critical add-ons to the flight deck of an aircraft carrier will be refueling drones. The Navy has awarded in 2020 a contract for 7 MQ-25 Stingray Carrier-Based Aerial-Refueling Systems (CBARS) [108] each of which is expected to increase the number of available strike-fighter of 6 per air-wing. Moreover, this technology will allow the wing to extend its range by 400 miles [109], while freeing more F/A-18E Super Hornet which were reconfigured to airborne tankers. Of course, these modifications will not come without investing on the ship itself. The Navy is planning to field these systems by 2024 while they install Unmanned Aviation Warfare Center rooms on some of the carriers.

Landing Craft Air Cushion (LCAC) are going to be an important technology upgrades for certain assets. LCAC provide over the horizon range with a contained radar cross section, they can be weaponized for achieving distributed lethality and they can carry up to 70 tons of equipment at about 40 knots. Some have suggest upgrades to the LCAC to achieve what is called the Fast Air Cushion Expeditionary Craft (FACEC) [110]. These FACEC could be used to carry up to 45 tons at the “near helicopter” speed of 85 to 100 knot for a range of about 200 miles, the open deck will provide space for modular weapons to be installed to provide for distributed lethality. The role of LCAC, or FACEC, is to provide a fast transport to shore to start a localized counter A2/AD position supporting naval operations and force projection [111].

vii. Unmanned Systems

Unmanned systems have seen an increase in relevance in the past decades becoming more relevant in Armed Forces. UAVs are extremely integrated in the Air Force for both ISR uses and precision striking. The Navy is also looking at increasing its unmanned capabilities: in September 2020 Secretary Esper outlined the vision for the 2045 fleet, which includes 140 to 240 unmanned and optionally maned vehicles [112]. According to the CSBA 2016 Fleet Architecture Study [24] unmanned asset should include both Extra-Large Displacement USV (XLUSV) like DARPA’s Sea Hunter – today considered a Medium-displacement USV – displacing more than 100 tons, and Common Unmanned Surface Vehicles (CUSV) about the size of a RHIB. XLUSV will be required to operate autonomously for most of the time with the option of being manned if needed. They will be used to support ISR functions, electronic warfare, mine countermeasure and ASW. According to DARPA platforms like the sea hunter will carry different payloads depending on the mission: from small UAVs to mines from a towed sonar to EM decoy equipment

[113]. The relatively small radar cross section will enable to employ these assets in contested scenarios. According to Clark [114], the US should look forward investing in a corvette-like USV, with a displacement of about 2000 tons. This will allow enough space on the ship to evolve over time while technologies mature, and it will give the US a sizable platform capable of navigating blue waters. A 2000-ton class ship will allow the USN to employ it in different scenarios, including contested one, controlling it and arming it with modern warfare systems. According to Clark, having the Sea Hunter at sea already performing complex tasks in autonomy will bring a positive impact to future unmanned technologies needed like autonomous decision making and persistent communications with the asset.

A.2 Underwater Vessels

In opposition to surface ships, the underwater world offers some unique advantages. Underwater assets have been capable to lunch long range standoff weapons since the middle of the cold war but today their stealth capabilities are even more relevant. Surface vessels can be spot from space while submarines, when navigating underwater, are invisible. They can penetrate A2/AD bubbles from below given the low diffusion of underwater protection systems. Nevertheless, navies must account for an increase and proliferation of anti-submarine technologies, specifically in littoral areas where big submarines might be too noisy to navigate. Hydrophones systems and Transformational Reliable Acoustic Path Systems (TRAPS) are getting more diffused among countries interested in keeping submarines at distance [115]. Some have suggested that the security environment shift will lead toward a less prominent role of submarines in direct operations, not considering the deployment of standoff weapons, toward a more C2 approach. In this, submarines will coordinate smaller, and stealthier, assets able to penetrate underwater defenses. It is expected that submarines will be able to launch UAV for EW to support aerial operations, to deploy large UUV which can themselves deploy smart UUV mines to engage underwater targets. New systems are also thought to achieve passive undersea surveillance [24]. Despite different defense systems it is always important to remember that near-silent modern submarines are hard to counter, any fleet approaching an area where enemy submarines are operating must take extensive precautions. This is even truer when a fleet is approaching coastal waters which are noisier than blue waters and that can greatly affect any detection capability.

i. Warfare Subsystem

Torpedoes are usually short-range weapons; due to the difficulty in detecting and countering them they can inflict substantial damage to any ship. Torpedoes have been updated in the years with more range and better sensors, but the concept have remained the same. The only technology disruptors that have been studied extensively are supercavitating torpedoes. These allegedly can reach a speed of 200 knots, compared to the 55 of the MK-48, but they are extremely noisy, and they have problems in steering their course to avoid disrupting the cavitation bubble.

Mines are also evolving. Smart mines are becoming more common as they could be used to hit specific target. Mines are also becoming smaller with moving capabilities in specific coastal areas for up to 800 nm. Controlled smart mine fields can automatically lineup in battle formation and can adapt their behavior depending on the sensed ships around them

[116]. Investments on mines are also focusing on their stealthiness with some manufacturer making them of irregular shape to confound sonars by camouflaging with the sea floor. Smart mines are also becoming more resilient to sweeping by data fusion intelligent fuses that can sense when someone is purposely trying to trigger the sensor [117].

Standoff weapons are considered pivotal for underwater systems, as such, there is an increase interest in having a large proportion of these safely stored in a submarine far from enemies' eyes. A clear expression of how important standoff weapons are is the Virginia Payload Module (VPM) which adds 4 7-slots vertical tubes for TLAMs, increasing the total number of deployable missiles in a Virginia class submarine from 12 to 40. It is important to mention that these modules will also be deep and large enough to carry hypersonic missiles, which might be why the USN decided to heavily invest on the Virginia class, reaching the cost of B$3.5per ship [118].

ii. Sensor & Processing Subsystem

Submarines most of the time use their sonar in passive mode to avoid being detected. Investments on sonars are leading toward better sensors that can identify underwater disturbances more precisely to avoid accidents like the 2009 one when two nuclear submarines, both using only their passive sonars, collided [119]. One of the areas in which countries are investing is the multistatic sonar. The idea behind it is that one active sonar produces the ping and multiple sensors in the water listen to the sound waves bouncing off enemy submarines. In this method dozens to hundred sensors are scattered to increase the chances of detecting the ping and therefore recognizing what kind of submarine is present and its direction and speed. This technology is really beneficial in multinational collaborations as different nations (e.g. NATO allies) can contribute with small arrays of sensors to create a vast underwater net. This multistatic sonar could be also paired with new development in AI which is now being used to support UUV and submarines in identifying diesel electric submarines ignoring the clutter of the coastal environment [120].

iii. Communication Subsystem

Most solutions for submarines communications are based on buoys which allow communications only in one direction. In situations of need the submarine can of course surface to periscope depth to allow communication via the antennas located in the mast, but in contested areas this can be risky. For these reasons, the US is investing in a laser technology communication system which will rely on the submarine communicating to a UUV at short distance which will relay to the Milstar satellite network. In this concept a fleet of small UUV stored in several underwater garages will be used to bridge the communication between the submarine and the satellites [121]. A different two-way communication system is what is being tested by the German Navy with the Callisto tethered buoy system. By using this system, the U212A class will be able to communicate two ways in short-range or one way in long range. The Callisto system allows link 11 and 16 on top of UHF sitcom and Battle Force emails [104].

iv. Structure & Hydrodynamics

A trend that has passed from ships to submarines is flexibility. There is a need to be able to tailor what is on board to specific missions; this can be achieved by reconfiguring some of the spaces inside the submarine. Flexibility in Virginia-class submarines is given by the reconfigurable torpedo room that can accommodate many special operations forces and all their equipment for prolonged deployments, as well as future off-board payloads. The VPM we have described before augment this capability [105].

A second aspect is the shape and material of the hull to reduce noise. Heavy investments have been done to reduce the noise level of nuclear submarine by employing anechoic coatings and isolated deck structures. Further investments are now focused on a new biomimetic propulsor which by seamlessly blending with the submarine structure, and by removing rotating parts, should further decrease noise level [121].

v. Propulsion Subsystem

All the submarines in the US Navy have a nuclear propulsion system. This allows them to have an unlimited underwater time, constrained only by food supplies on board. Speed of course is one of the advantages of nuclear propulsion, allowing submarines to maintain a constant 35 mph speed when submerged. Noise wise, the hydraulics of the nuclear reactor needs to pump coolant inside generating noise. Diesel is a valid alternative as it comes at a much cheaper cost compared to nuclear submarines. Nevertheless, unless they run on batteries, which typically limits their speed and it's limited for few hours, diesel submarines are noisier and have a shorter endurance. As a compromise many modern navies have invested in Air Independent Propulsion (AIP) systems which provides longer endurances – up to 4 weeks underwater depending on the technology used for the AIP – quieter hulls as AIP engines are virtually silent, and they come at a quarter of the price of a nuclear reactor [122]. Nuclear submarines have still the advantage on longer endurance missions, but a navy thinking about patrolling a specific area might want to invest on AIP submarines as it could create a much larger fleet with same resources of fewer nuclear vessels.

New technologies are focusing now on biomimetic propulsion systems. These are expected to reduce the noise of the submarine by removing rotating parts as the drive shaft and the blades of the propellers. These system are being tested on small USV but they could be considered also for bigger submarines in the future [123].

vi. Unmanned Systems

Underwater Unmanned Vehicles (UUV) have seen a huge push in the latest years with assets going from over 30ft in length to others being just few inches. Extra Large UUV (XLUUV) can be employed in long patrol missions having on board diesel generators which can have them running for more than 6 months and 1500nm. These can use both active and passive sonars to support underwater surveillance operation without risking the use of active sonars on a manned platform. Some of the Large UUV (LUUV), up to 30ft [24], could be deployed from one of the tubes of the VPM on Virginia-class submarines in place of the seven Tomahawks usually loaded [124]. Other LUUV that are receiving more attentions are gliders. These can be used for long duration surveillance using as propulsion system waves and underwater currents and small motors recharged by solar panel on their surface structure. Gliders can carry small passive sonars, but they can communicate as they have an above water component. Medium and Small UUV have the size of a Mk-48 and Mk-18 torpedo, respectively. Both can be used either for surveillance or for carrying weapons. Depending on the specific setting and on the power supply they can be used within a range up to 1000 nm. Their reduced noise and small dimensions make them ideal in A2/AD contexts where underwater sensors are placed around the area. Finally, investments in Micro UUV (smaller than 6 inches) are leading toward the use of small swarms of UUV for attacking infrastructures, jamming and decoying sonars or monitoring smaller areas [125].

APPENDIX B. Bootstrapping Analysis

When dealing with a stochastic simulation with stochastic responses, it is important to perform a certain number of repetitions to find the confidence intervals for the mean of the responses as well as to determine how the response is distributed. For each response, several repetitions must be determined, and the maximum number of repetitions needed for any important response must be run for each of the Design of Experiments cases. A technique called Bootstrapping was used to determine the number of repetitions needed in the many experiments of this thesis. Through Bootstrapping, it is possible to estimate how the variance will change given the number of repetitions of the experiment. Numbers provided as examples here were taken from the experiment represented in CHAPTER 10 10.

Bootstrapping can be summarized in the following four steps:

6. A large sample of the baseline case is created
    a. In our case we used 800 repetitions for each deterrence strategy keeping the basic NATO fleet and all Asyr assets at their nominal level
7. The large sample is resampled with replacement to get many groups of samples with varying size
    a. In our case we created groups of multiple of 5, from 5 to 100
8. For each group, the mean is calculated
9. By looking at the distribution of the mean the standard deviation of the mean is calculated

10. Even if not always needed, it was decided to normalize the standard deviation by the mean to be able to study the *coefficient of variation* c_V and its reciprocal, the *signal to noise ratio* SNR

SNR and c_V are the key to understanding how many repetitions are required for an accuracy level. For a study using a computationally expensive simulation, it is important to reduce the number of repetitions as much as possible without sacrificing on the SNR, i.e., increase statistical error to unacceptable levels. Having a low *coefficient of variation* means being able to trust more the mean results of the experiment.

Looking at Figure 61, it is possible to see how the variable “Number of Ally Killed” has two distinct kinks at different Number of Groups: one around 40 and a second one around 85. Therefore, in the effort to minimize the *coefficient of variation*, it is desired to have at least 40 repetitions of the experiment. In fact, in this case, 40 represents the sweet sport in which we get accurate results at a manageable computational cost.

In the end, it is worth mentioning that while reaching the kink is advisable to get the optimal combination between quality of results and computational efforts, this is not always needed. In fact, if the value of the coefficient of variation is sufficiently low, results are acceptable even if the repetition number hasn’t reached the kink yet.

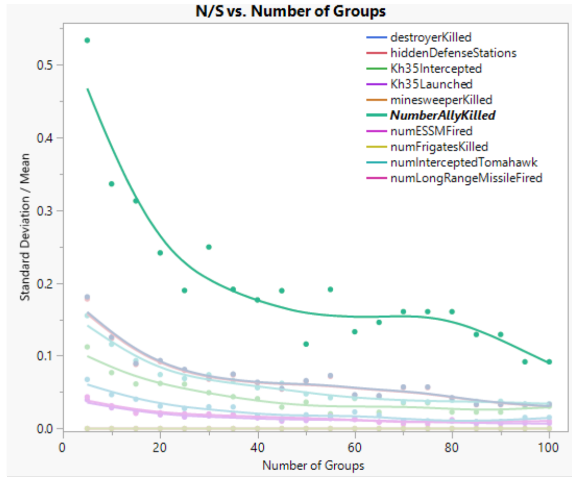
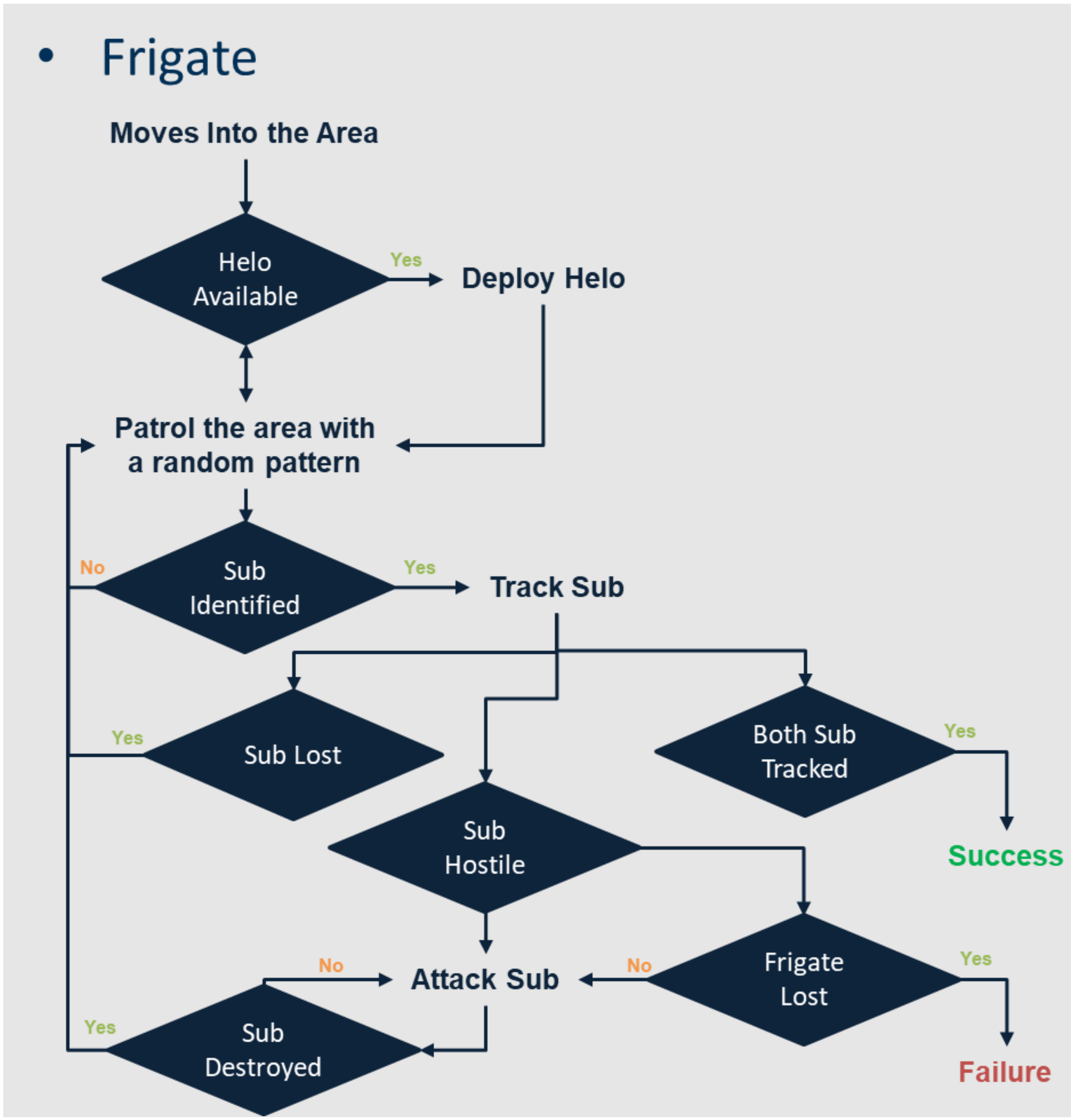
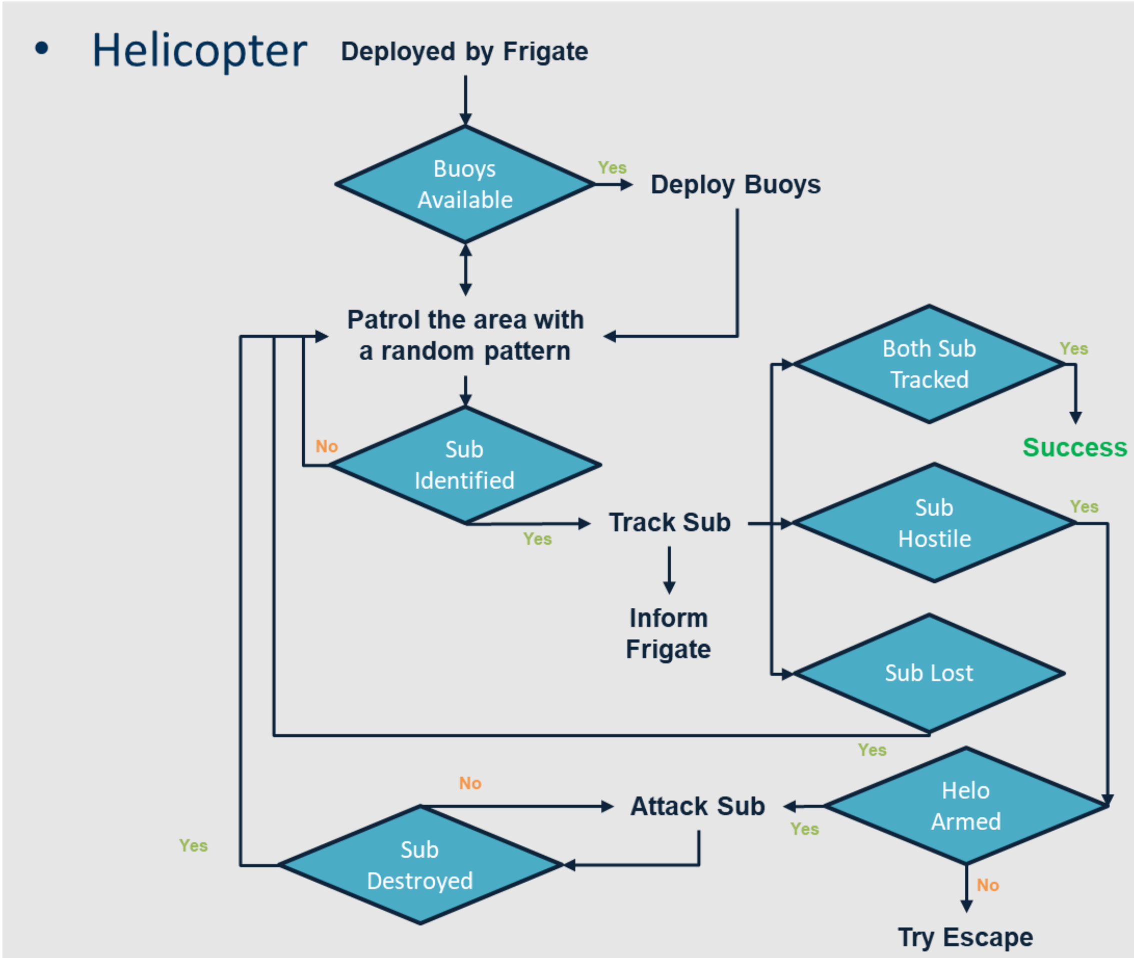
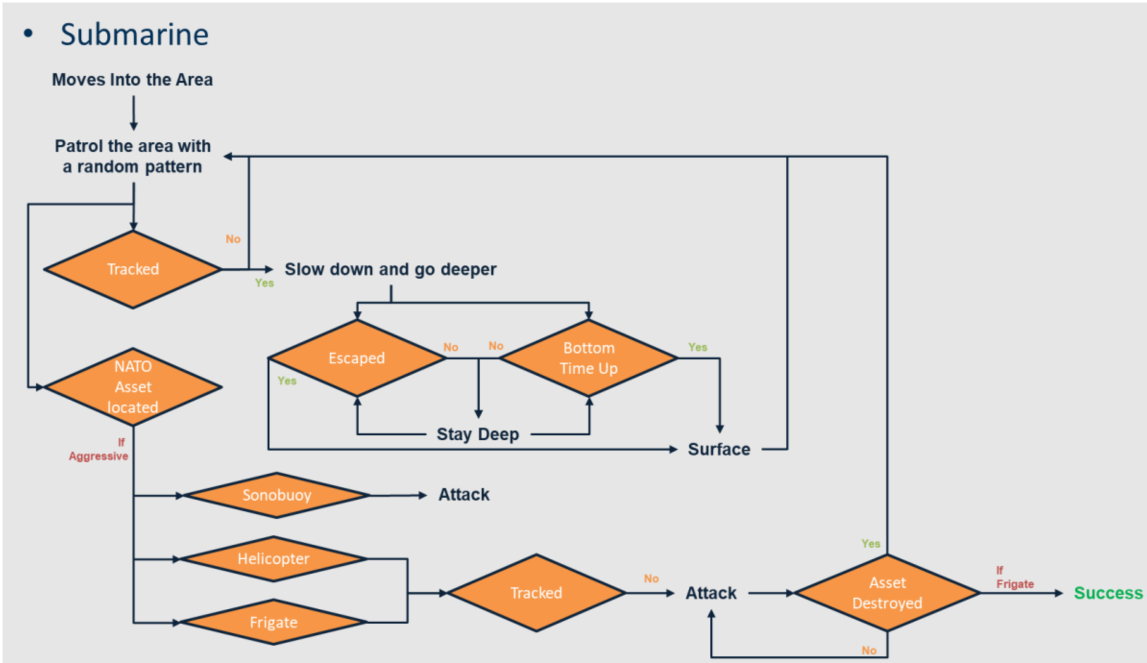


Figure 61: Representation of the two kinks used in Bootstrapping to select the number of repetitions in the simulation

APPENDIX C. Chapters 6 to 9 assets' behavior logical diagrams

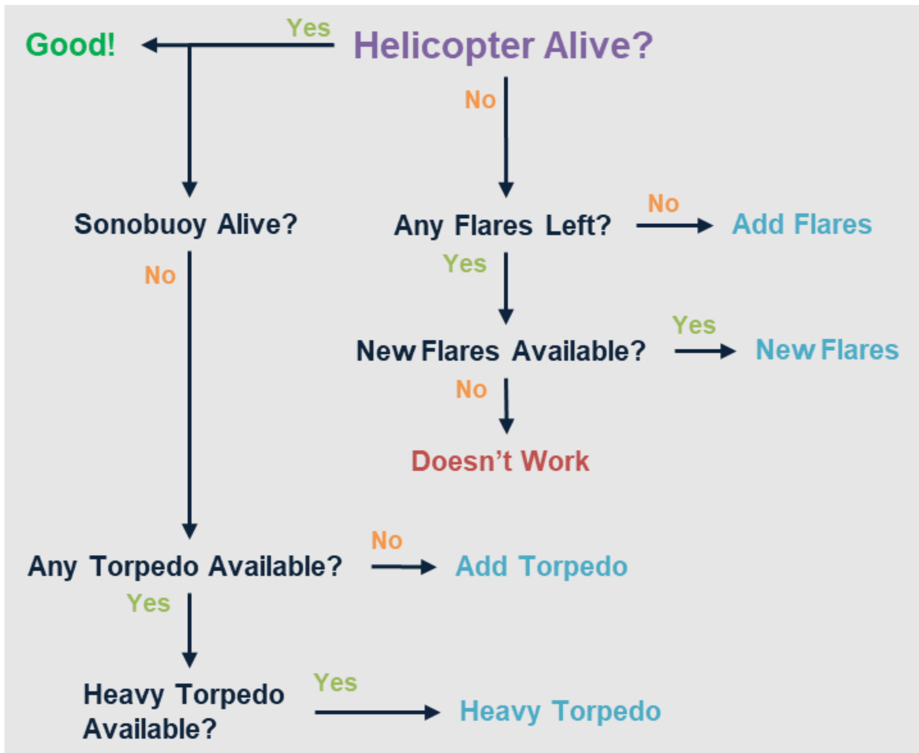


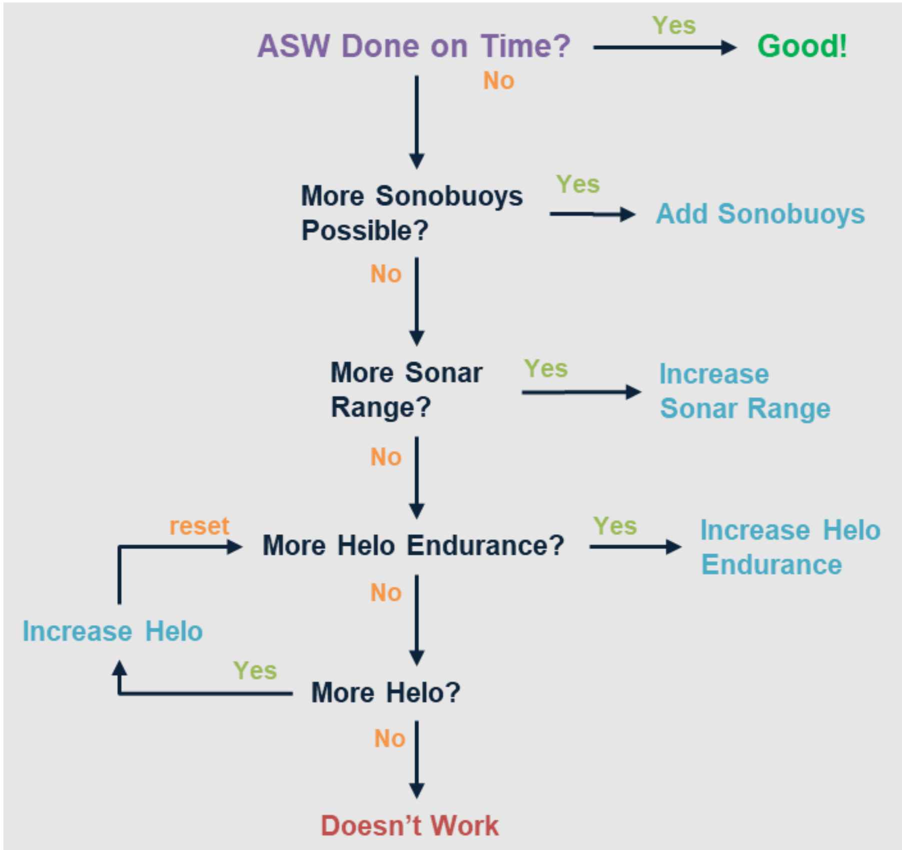

• Helicopter
Deployed by Frigate
Buoys Available
Yes
Deploy Buoys
Patrol the area with a random pattern
No
Sub Identified
Yes
Track Sub
Inform Frigate
Both Sub Tracked
Yes
Success
Sub Hostile
Yes
Sub Lost
Yes
Helo Armed
Yes
Attack Sub
No
Sub Destroyed
Yes
No
Try Escape



Submarine
Moves Into the Area
Patrol the area with a random pattern
Tracked
No
Yes
Slow down and go deeper
Escaped
No
No
Bottom Time Up
Yes
Yes
Stay Deep
Surface
NATO Asset located
If Aggressive
Sonobuoy
Attack
Helicopter
Frigate
Tracked
No
Attack
Yes
Asset Destroyed
If Frigate
Success
No

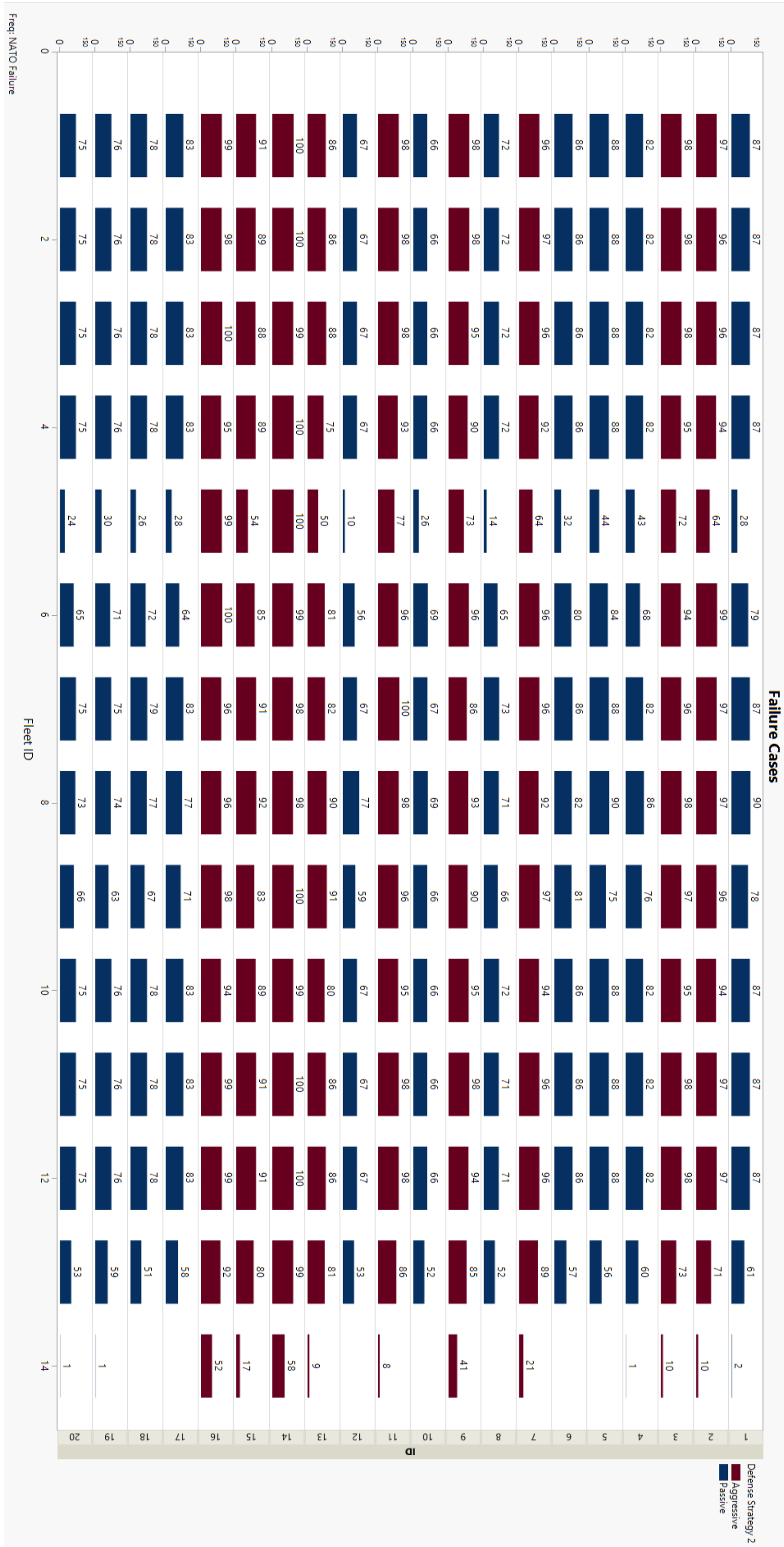
APPENDIX D. Decision Trees for the Iterative Algorithm in Chapter 7



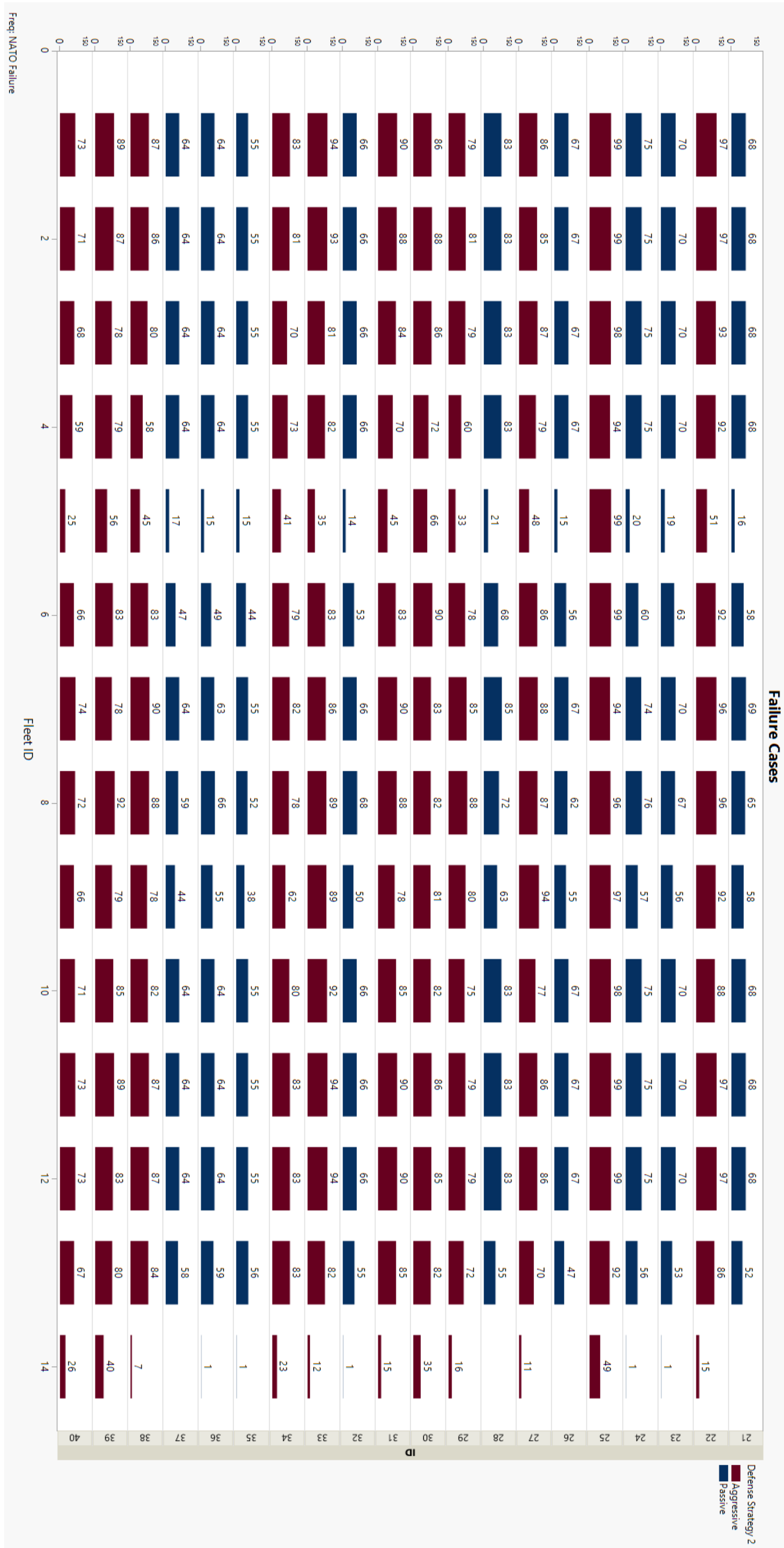

ASW Done on Time?
Yes
Good!
No
More Sonobuoys Possible?
Yes
Add Sonobuoys
No
More Sonar Range?
Yes
Increase Sonar Range
No
reset
More Helo Endurance?
Yes
Increase Helo Endurance
No
Increase Helo
Yes
More Helo?
No
Doesn't Work

APPENDIX E. Complete Results from Experiment in Chapter 8

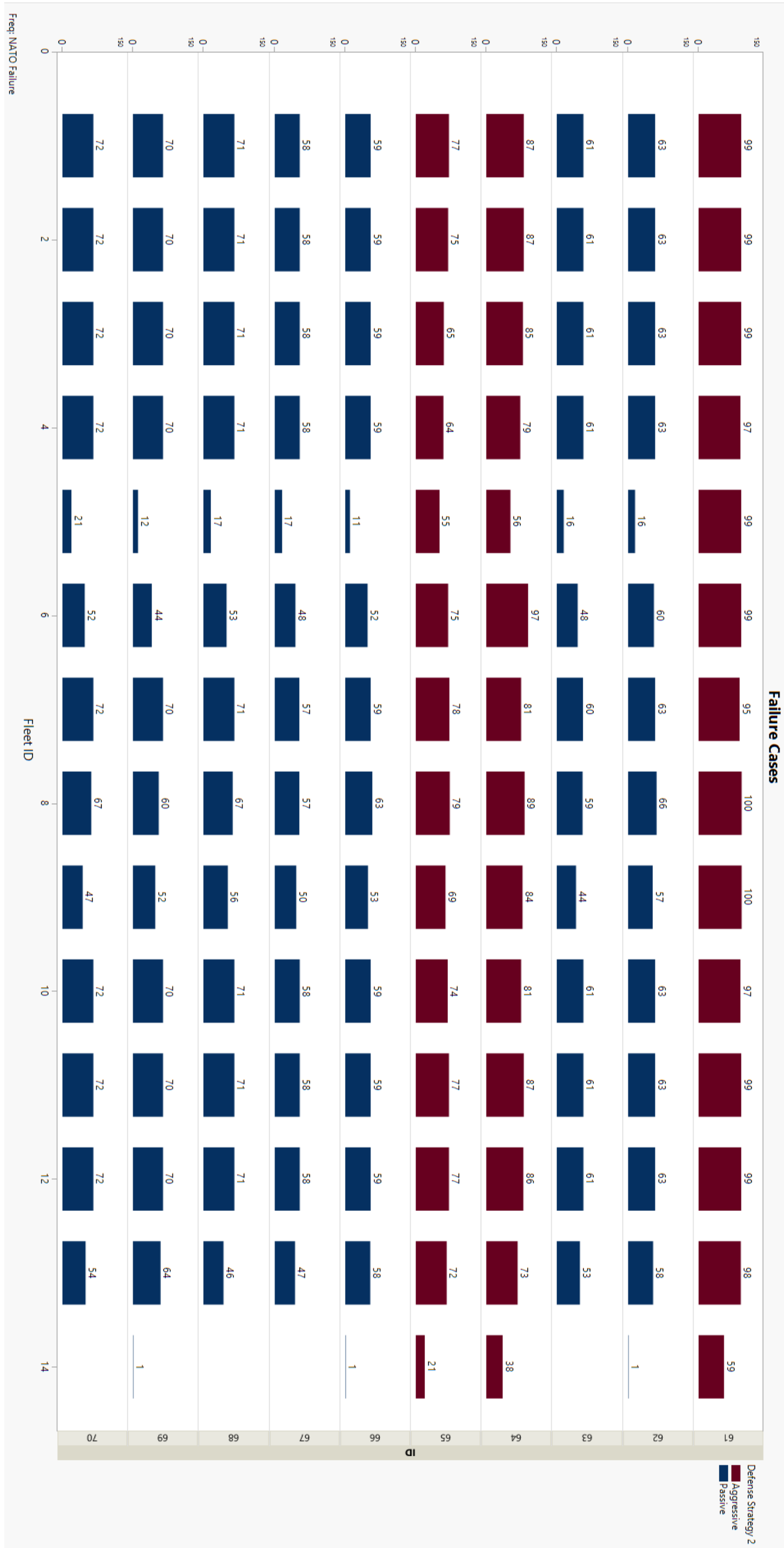
The following figures are the full results from the JANUS simulations conducted in Chapter 8. They show the failure rate of cases per fleet tested per scenario. The maximum value in the y-axis is 100 as there are 100 repetitions each fleet could fail. The minimum value is zero, this means that all the simulations in that fleet in that scenario were successful.



Failure Cases
Fleet ID
Freq: NATO Failure
ID
Defense Strategy 2
Aggressive
Passive


Failure Cases
Fleet ID
ID
Freq: NATO Failure
Defense Strategy 2
Aggressive
Passive


Failure Cases
Fleet ID
Freq: NATO Failure
ID
Defense Strategy 2
Aggressive
Passive


Failure Cases
Fleet ID
ID
Freq: NATO Failure
Defense Strategy 2
Aggressive
Passive

APPENDIX F. Comprehensive List of Assumption for Chapter 10 Experiment

The purpose of this appendix is to provide a reference for all the different assumptions that were taken throughout the experiment in Chapter 10. Assumptions are divided per faction.

A. Asyr

A1. Asyr has available only a limited number of aircraft (up to 20) which can be equipped with anti-ship missile Kh-35.

A1I. Each aircraft can be equipped with 1 to 4 anti-ship missiles.

A1II. Fighters are all launched from the only air force base in the Northwest side of the island.

A1III. Aircrafts' behavior is connected to the one of the defense stations. In the Saturation All-out deterrence strategy aircrafts and defense stations act synchronously coordinated by the C2 center.

A1IV. If the C2 center is destroyed then aircraft will have to act on their own. They will be able to attack, but they won't be able to coordinate with other defense stations.

A1V. If the C2 center is destroyed before aircraft are launched they won't be launched.

A2. Asyr has one main command and control (C2) center located in the Northwest side of the island.

- A2I. If the C2 center is destroyed by the NATO fleet Asyr cannot coordinate the attacks anymore, each defense station will act autonomously.
- A2II. The C2 center is defended by fixed 2 Surface-to-Air Missiles (SAM) stations.
- A2III. Each SAM station has 10 short range missiles ready to launch

A3. Asyr has a variable number of defense stations (0 to 10) which are made of Russian truck Bal-E.

- A3I. Each truck has a variable number of missiles (4 to 10) Kh-35 ready to launch.
- A3II. Each truck can reload its launcher a finite amount of time (0 to 3), each reload has the same number of missiles of what was originally ready to launch.
- A3III. Reloading time is 15 minutes.
- A3IV. Each truck is defended by a ground unit with shoulder fired Surface-to-Air Missiles, with up to 10 MANPADS.
- A3V. Each truck is assumed to be in a fixed position. At the beginning of the simulation NATO fleet will receive intelligence information. Each truck has a variable chance of being known to the incoming NATO fleet.
- A3VI. If a truck position is known, then that truck can be fired upon by the NATO fleet.

A3VII. After a station fires its first missile its position will be known to the NATO fleet.

A3VIII. Each station has a detection and engagement range of up to 50nm.

A4. The missile used by both defense stations and planes is the anti-ship missile Kh-35.

A4I. Three quality levels are defined for the Kh-35: High/Medium/Low

A4II. The Kh-35 can mount on board different sensor suites (edge tracker, IR, Optical, or a combination of the three). The better the sensor, the higher the P_Kill.

1. P_kill: Against Frigate, High: 0.9, Medium:0.7, Low:0.5
2. Destroyer: 0.8/0.6/0.4
3. TG: 0.6/0.5/0.4
4. MS: 1/0.8/0.6

A4III. Each version of Kh-35 has a different survivability against NATO defense systems.

1. P_survivability: High: 0.3 against Long range/0.6 against short range/ 0.95 against CIWS
2. Medium: 0.2/0.5/0.925
3. Low: 0.1/0.4/0.9

A4IV. Each version of Kh-35 has different maneuverability and targeting abilities

1. High: Once the seeker is activated the missile can change its course to intercept the target. If the target has been destroyed

by a previous missile it can switch to a new target. Large explosion blast (150m).

2. Medium: Once the seeker is activated the missile can change its course to intercept the target. Medium explosion blast (125m).
3. Low: No steering capability, small explosion blast (100m).

A5. Asyr has four different tactics that can use against the incoming NATO fleet: Passive, Persuasive, Saturation and All-out

A5I. If Asyr has a Passive strategy no defense station will activate its missiles, nor will planes be launched. Asyr, will let the NATO fleet approach without shooting. This is not used in the scenario, but just as a debug function.

A5II. In the Persuasive Strategy each defense station will shoot at the NATO fleet as it enters in its range. When shooting each defense station will deplete its available load, equally distributing the missiles across the incoming fleet, and starting the reloading process. Aircrafts will be deployed and will patrol the area where the NATO fleet is expected to arrive. When in range, aircrafts will launch their Kh-35 missiles equally distributing them across the incoming fleet.

A5III. In the Saturation Strategy all defense stations will send out a message to the C2 center and will wait for an order to attack all at the same time when the fleet is range of all defense stations. For all the defense stations the target is the High Value Unit of the NATO fleet. If the HVU is destroyed, but the defense stations have still missiles they will switch to other targets firing at will. Aircrafts will be deployed and will patrol the area where the NATO fleet is expected to arrive. When in range fighters will launch their Kh-35 missiles equally distributing them across the incoming fleet. Aircrafts will delay their attack to try to neutralize what is left in the fleet after the first wave from the defense stations.

A5IV. In the All-Out Strategy all defense stations will send out a message to the C2 center when they detect the fleet. DS will wait for an order to attack all at the same time when the fleet is range of all defense stations. For all the defense stations the target is the High Value Unit (HVU) of the NATO fleet. Aircrafts will stay in the back and coordinate with the defense stations, via the C2 center, to attack at the same time. Aircrafts will attack all the ships in the fleet to saturate their defense trying to protect the missiles from the defense stations headed to the HVU. The C2 center will distribute shooting orders in such a way that missiles will reach the target at the same time.

A5V. The C2 center is considered a high value target, as such it is defended accordingly. If it gets destroyed assets lose the ability to communicate unless close to each other (fighters still have coms among them, but defense stations are isolated). If it gets destroyed all strategies are downgraded to the Persuasive one as coordination is not possible. If the C2 center gets destroyed before fighters are launched, fighters won't be launched.

N. NATO

N1. Purpose

N1I. The portion of the scenario simulated in this use case starts once the fleet has reached the archipelago and it is about to enter Asyr's water.

N1II. The goal of the fleet for this portion of the scenario is to achieve control over the waters of the channel. To do so it must:

1. Survive any attack from known and unknown threats, both land and air based.
2. Neutralize any defense position known
3. Neutralize the enemy C2 center (optional)

N2. Composition

N2I. The NATO fleet is made of the following ships:

1. 1 CV (ITS Cavour type) – HVU
2. 2LHD (FS Mistral) – HVU
3. 1 DD (USS Arleigh Burke)
4. 7 FF (HNLMS De Zeven Provinciën / FREMM)
5. 1 AOR (USS Henry J. Kaiser) – HVU

N2II. To reduce computational efforts the CV, the 2 LHD and the AOR are modelled as a single HVU

N2III. No aerial component is modelled for this segment of the mission

N2IV. The fleet is assumed to be performing a single task – Suppression of Enemy Defenses – without having to consider other tasks like ASW, AAW, MCM, and NEO

N3. RoE – the Rules of Engagement defined in the scenario are that the NATO fleet can fire only if fired upon.

N3I. For this simulation it is assumed that the fleet was already fired upon and therefore has the liberty to fire as deemed appropriate against military targets.

N4. Armaments and Technologies Under Study

N4I. Hyper Velocity Projectiles (HVP) could be used instead of conventional ammunitions to increase the P_Kill of common CIWS and Naval guns

N4II. New ESSM Sparrow can be loaded in the VLS to increase the P_Kill of these short-range missiles. These can be also quad-packed to increase the number.

N4III. New SM-2/3/6 can be loaded for an increased P_Kill on these long-range missiles

N4IV. Additional naval guns can be mounted (up to 3 per ship) to increase the resiliency of each ship. If those are mounted they should be of the lowest caliber available on the ship. In the simulation they will provide additional independent fire points on each ship.

N4V. Additional VLS Blocks can be added (up to 5 per ship) to carry more missiles. Each block adds additional 8 tubes to the ship. The ratio of the missiles inside is the same of the main VLS.

N4VI. Radar range can be increased to provide for early detection of enemy assets. Asyr defense stations that were not detected by intelligence won't be detected until they shoot their first missile.

N4VII. The VLS (and all additional blocks) can be configured in 3 possible ways: Standard, Aggressive and Defensive. The Standard one has 25% Tomahawks, 15% ESSM, and 60% SM-2/3/6. The Aggressive one has 35% Tomahawks, 10% ESSM, and 55% SM-2/3/6. The Defensive has 15% Tomahawks, 20% ESSM, and 65% SM-2/3/6. ESSM slots are quad-packed.

N4VIII. If one technology is enabled it is applied to all the ships in the fleet that have that capability (adding one VLS block, will add 8 tubes to all ships in the fleet that have a VLS).

P. Prucy

P1. Prucy has no military forces, so while it is thought to be neutral by NATO it will allow in its territory some Asyr troops. These include a series of mobile trucks type Bal-E with Kh-35 missiles and a small number of troops for point defense.

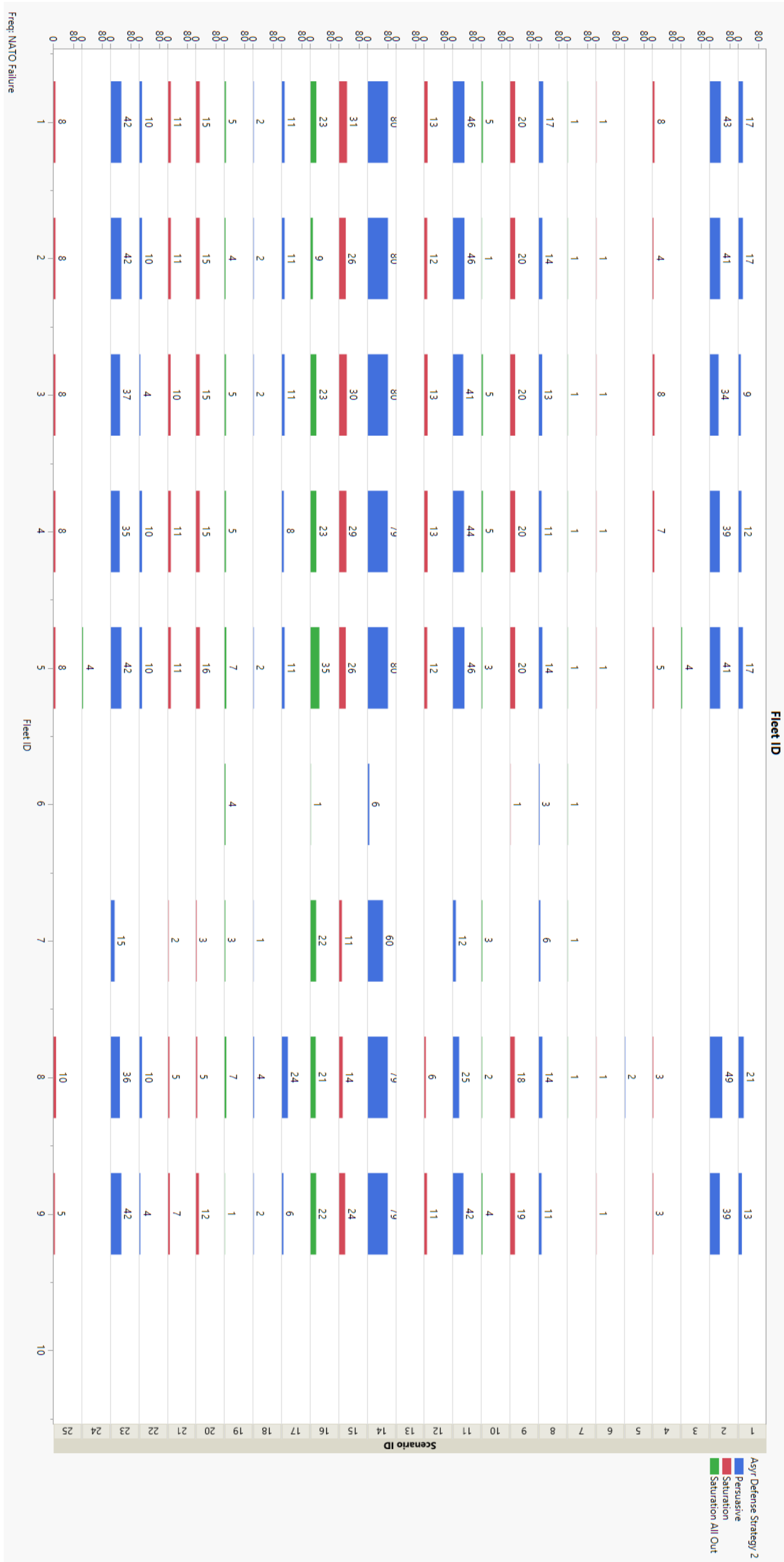
APPENDIX G: Full Results from Chapter 10 - Part 3.1

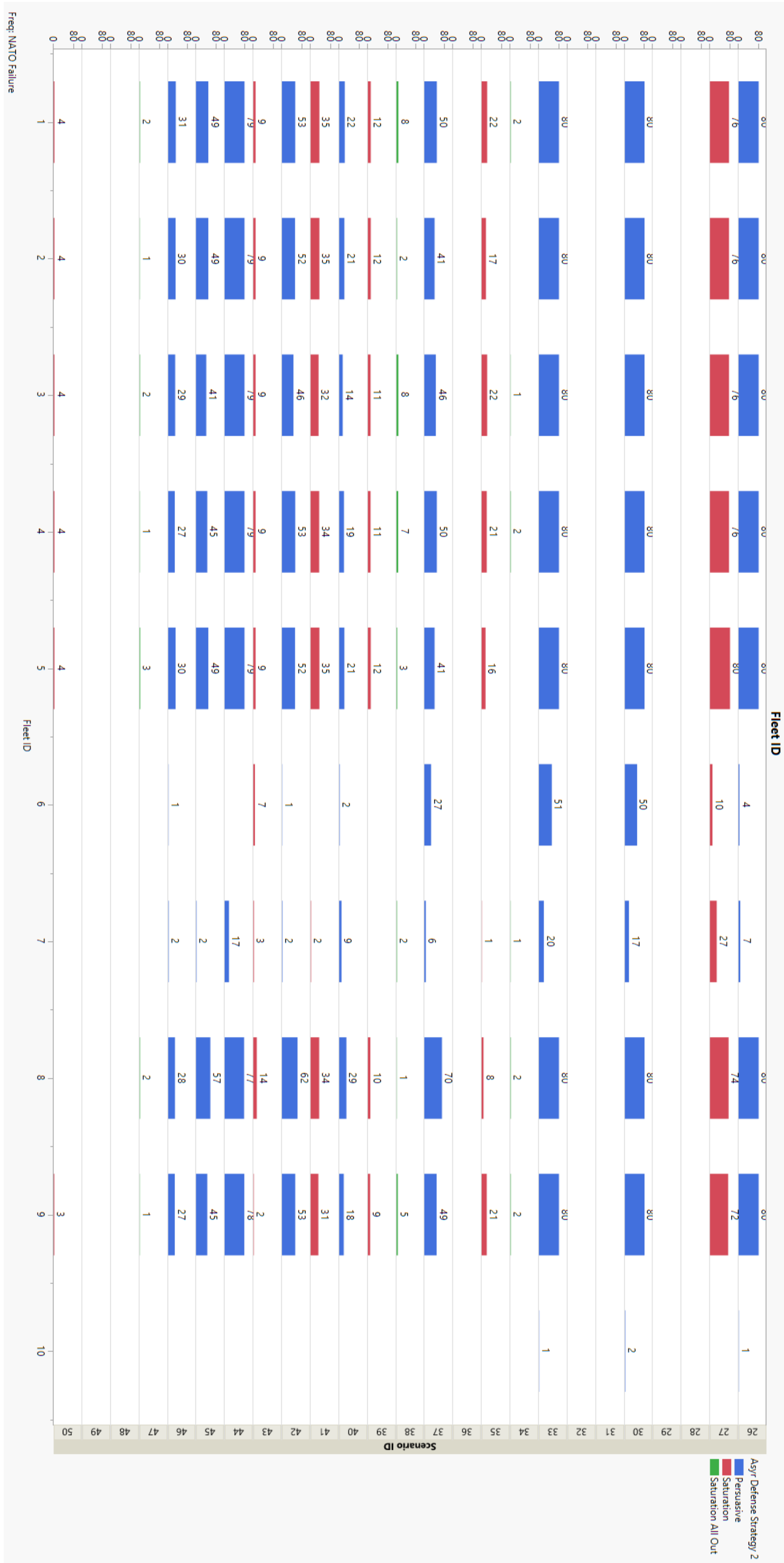
The following figures are the full results from the JANUS simulation of the 3rd part of the main experiment. The focus of this experiment is the verification of the positive monotony of the technology used in the fleets. Figures below show the failure rate of cases per fleet tested per scenario. The maximum value in the y-axis is 80 as there are 80 repetitions each fleet could fail. The minimum value is zero (which is not showed in columns), this means that all the simulations in that fleet in that scenario were successful. Information regarding how each fleet was generated can be found in Table 23. In this experiment only one technology at the time was maximized, so each fleet shows the benefit or the issues of having a specific technology onboard.

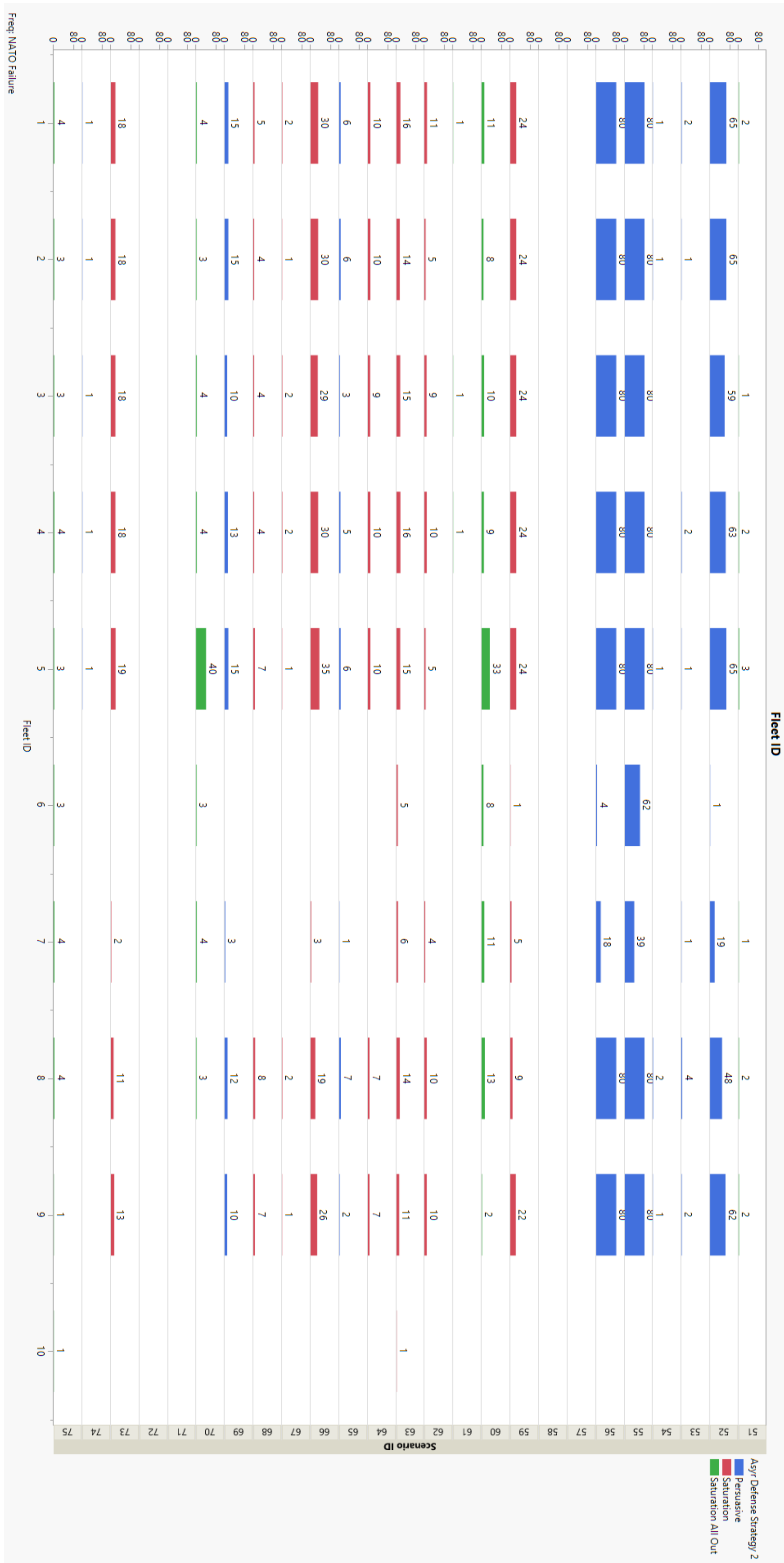
Table 24: Description of the technologies modified in each of the fleet used in the Part 3 of the main experiment

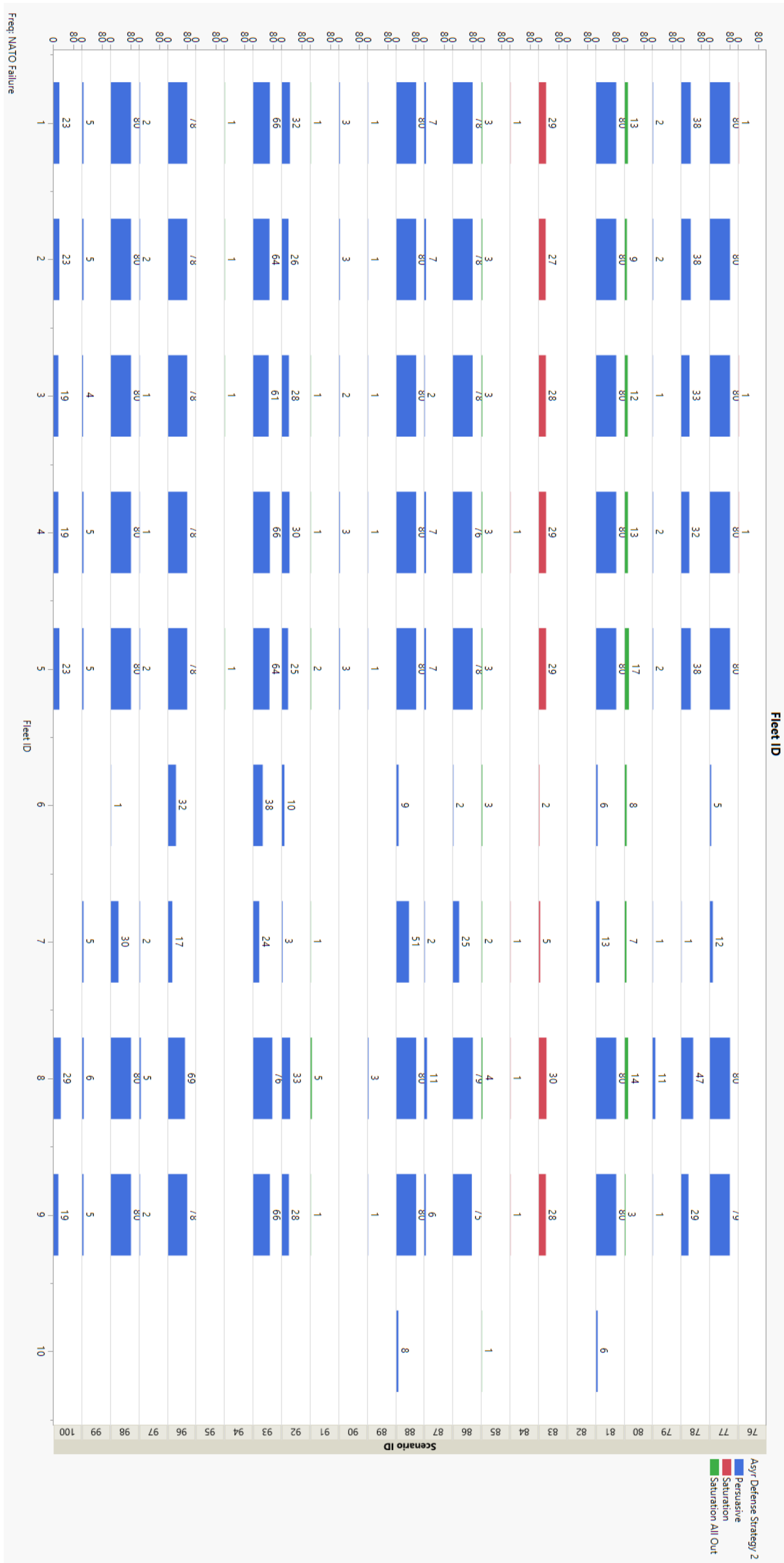
FLEET	MODIFIED PARAMETER
1	Basic Fleet as in main experiment part 1
2	Hyper-Velocity Projectiles used
3	New Short-Range missiles used
4	New Long-Range missiles used

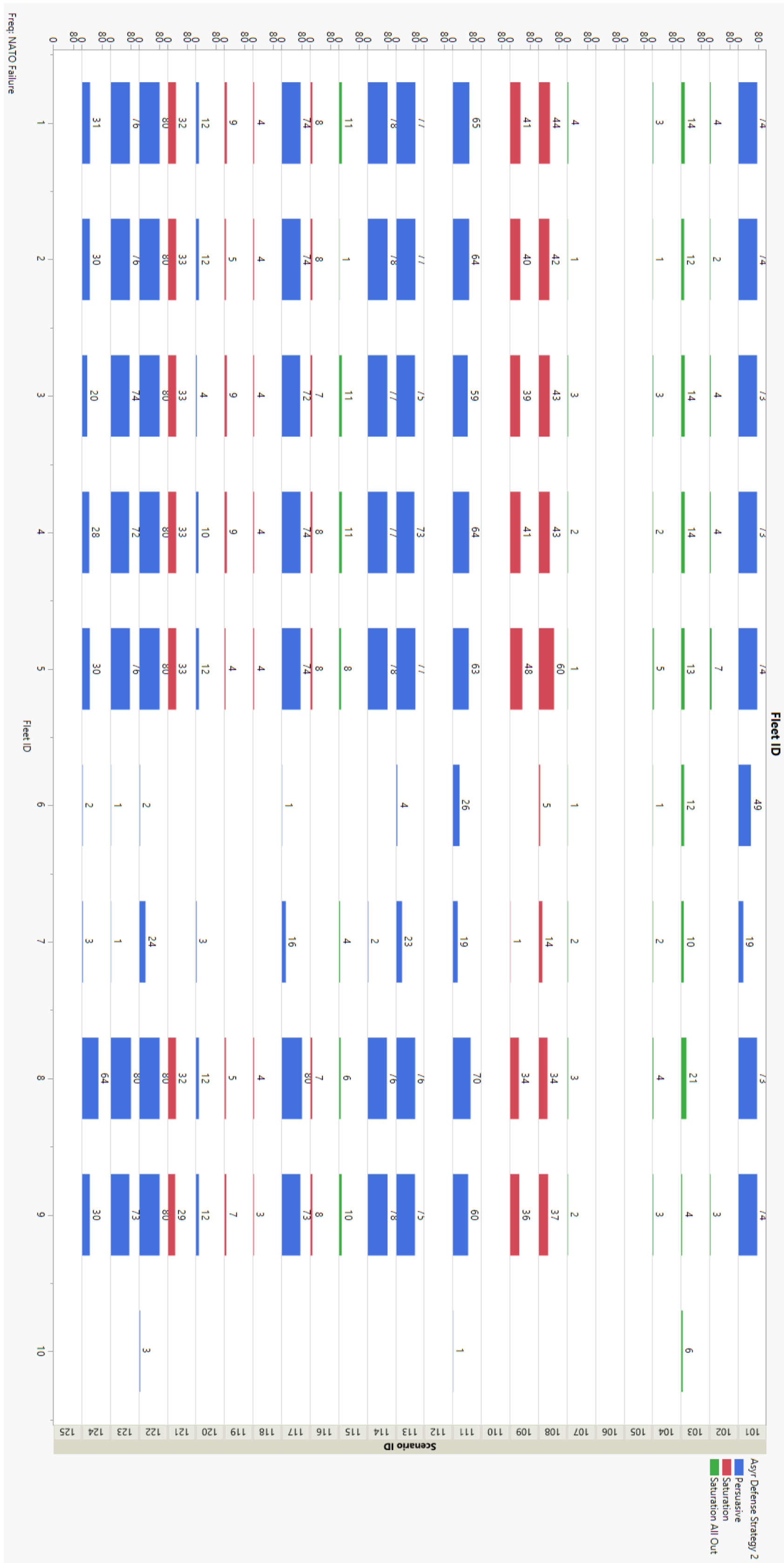
5	Naval Guns maximized
6	Hull Strength maximized
7	Radar Range maximized
8	VLS Fire Rate maximized
9	New TLAM used
10	All technologies enabled and maximized

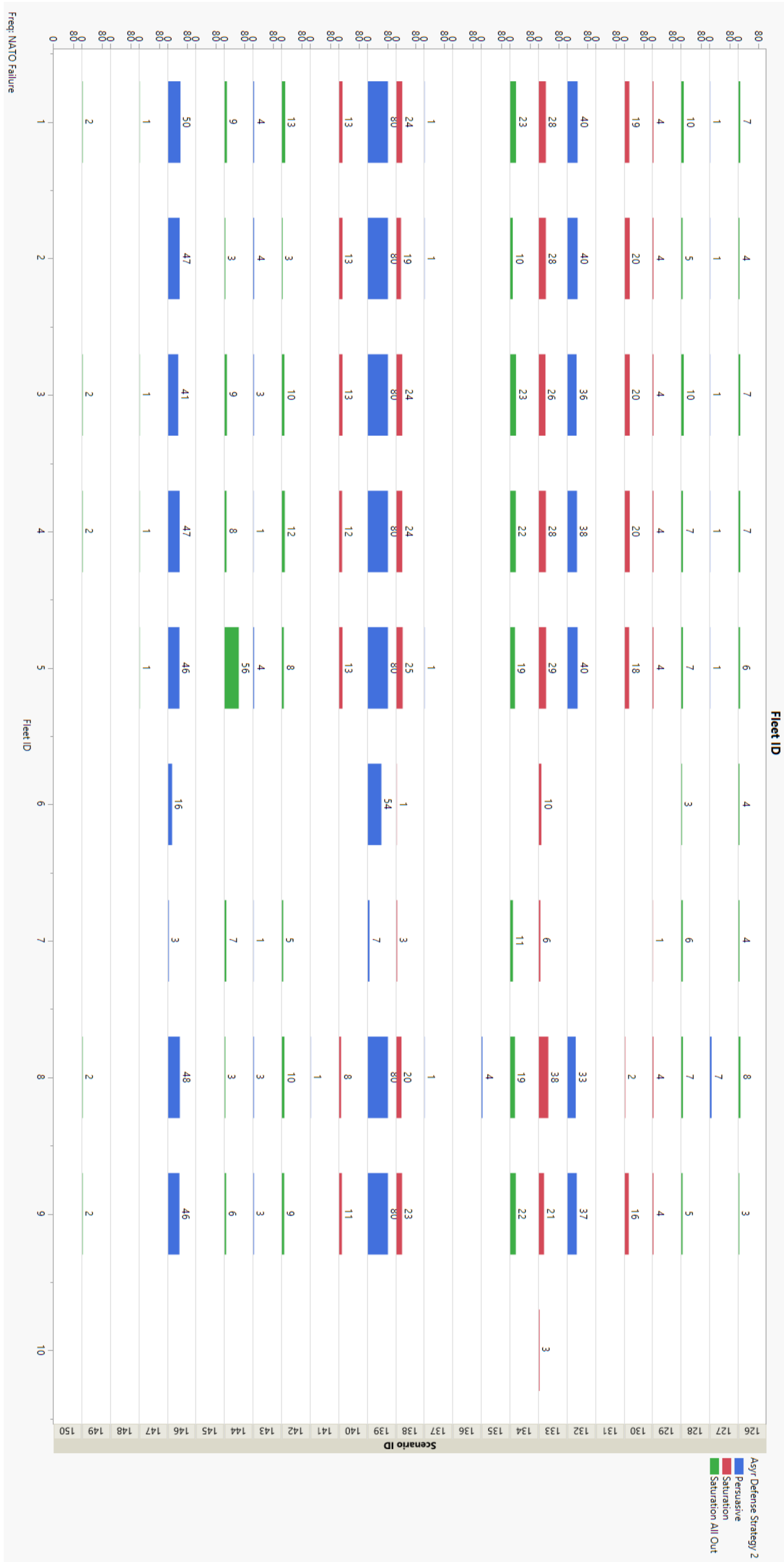

Fleet ID
Scenario ID
Freq: NATO Failure
Asyr Defense Strategy 2
Persuasive
Saturation
Saturation All Out


Fleet ID
Scenario ID
Freq: NATO Failure
Asyr Defense Strategy 2
Persuasive
Saturation
Saturation All Out


Fleet ID
Scenario ID
Freq: NATO Failure
Asyr Defense Strategy 2
Persuasive
Saturation
Saturation All Out


Fleet ID
Scenario ID
Freq: NATO Failure
Asyr Defense Strategy 2
Persuasive
Saturation
Saturation All Out


Fleet ID
Scenario ID
Freq: NATO Failure
Asyr Defense Strategy 2
Persuasive
Saturation
Saturation All Out


Fleet ID
Scenario ID
Freq: NATO Failure
Asyr Defense Strategy 2
Persuasive
Saturation
Saturation All Out

References

[1] J. A. Winnefeld, "Winter Is Coming," vol. 146, no. July, pp. 1–2, 2020.

[2] E. G. Kail, "Leading in a VUCA environment: V is for volatility," *Harv. Bus. Rev.*, pp. 1–5, 2011, [Online]. Available: https://hbr.org/2011/01/leading-effectively-in-a-vuca-1.

[3] S. G. Jones, "War by Proxy: Iran's Growing Footprint in the Middle East," *Cent. Strateg. Int. Stud.*, no. March, pp. 1–16, 2019.

[4] R. Malley, "The Unwanted Wars ," *Foreign Affairs*, 2019.

[5] F. Farhi, "Iranian Power Projection Strategy and Goals," no. April, pp. 1–12, 2017, [Online]. Available: https://csis-prod.s3.amazonaws.com/s3fs-public/publication/170421_Farhi_Iranian_Power_Projection.pdf?R.Ck9Gr6VPAvyWbG2JIPf3abDR201S1.

[6] Group Insikt, "Despite Infighting and Volatility , Iran Maintains Aggressive Cyber Operations Structure," 2020.

[7] "BBC - History - World Wars: The Fall of France." http://www.bbc.co.uk/history/worldwars/wwtwo/fall_france_01.shtml (accessed Nov. 29, 2020).

[8] J. Mattis, "U . S . Military Adapting in Complex Environment," pp. 1–5, 2017.

[9] M. Fitzsimmons, “The False Allure of Escalation Dominance,” *War Rocks*, pp. 1–12, 2017, [Online]. Available: https://warontherocks.com/2017/11false-allure-escalation-dominance.

[10] D. Mercier, “NATO’s Adaptation in an Age of Complexity.”

[11] E. G. Kail, “Leading Effectively in a VUCA Environment: A is for Ambiguity,” 2011. https://hbr.org/2011/01/leading-effectively-in-a-vuca-1 (accessed Sep. 22, 2020).

[12] M. Roberto, R. M. J. Bohmer, and A. C. Edmondson, “Facing Ambiguous Threats,” *Harv. Bus. Rev.*, vol. 11, 2006, Accessed: Sep. 22, 2020. [Online]. Available: https://hbr.org/2006/11/facing-ambiguous-threats.

[13] Center for Strategic and International Studies, “Competing in the Gray Zone | Center for Strategic and International Studies.” https://www.csis.org/features/competing-gray-zone (accessed Sep. 22, 2020).

[14] A. C. Kuchins, “Crimea One Year On: Where Does Russia Go Now?,” *Center for Strategic and International Studies*, 2015. https://www.csis.org/analysis/crimea-one-year-where-does-russia-go-now (accessed Sep. 22, 2020).

[15] Council on Foreign Relations, “China’s Martime Disputes,” 2019. .

[16] A. Chorn and M. M. Sato, “Maritime Gray Zone Tactics: the Argument for Reviewing The 1951 U.S.-Philippines Mutual Defense Treaty,” *New Perspect. Foreign Policy*, no. 18, pp. 3–9, 2019, [Online]. Available:

https://www.csis.org/maritime-gray-zone-tactics-argument-reviewing-1951-us-philippines-mutual-defense-treaty.

[17] G. B. Poling, "How Significant Is the New U.S. South China Sea Policy? ," 2020. https://www.csis.org/analysis/how-significant-new-us-south-china-sea-policy (accessed Sep. 22, 2020).

[18] R. O'Rourke, "Navy force structure and shipbuilding plans: Background and issues for congress," *Def. Econ. Recent Prog. Remain. Challenges*, pp. 149–268, 2018.

[19] Department of the Navy, "2016 Navy Force Structure Assessment (FSA)," 2016.

[20] D. B. Larter, "In a quest for 355 ships, US Navy leaders are unwilling to accept a hollow force," 2020. https://www.defensenews.com/digital-show-dailies/surface-navy-association/2020/01/13/in-a-quest-for-355-ships-us-navy-leaders-are-unwilling-to-accept-a-hollow-force/ (accessed Sep. 28, 2020).

[21] Naval History and Heritage Command, "US Ship Force Levels." https://www.history.navy.mil/research/histories/ship-histories/us-ship-force-levels.html#2000 (accessed Sep. 28, 2020).

[22] D. B. Larter and A. Mehta, "The Pentagon is eyeing a 500-ship Navy, documents reveal," 2020. https://www.defensenews.com/naval/2020/09/24/the-pentagon-is-eyeing-a-500-ship-navy-documents-reveal/?utm_source=clavis (accessed Sep. 28, 2020).

[23] MITRE, "Navy Future Fleet Platform Architecture Study," 2016, [Online].

Available: http://www.mccain.senate.gov/public/_cache/files/1a3e3a4e-6c97-42fb-bec5-a482cf4d4d85/mintre-navy-future-fleet-platform-architecture-study.pdf.

[24] B. Clark, P. HAYNES, B. MCGRATH, C. HOOPER, J. SLOMAN, and T. A. WALTON, "Restoring American Seapower, A New Fleet Architecture For the United States Navy," pp. 297–350, 2017.

[25] U. Navy, "Alternative Future Fleet Platform Architecture Study," 2016.

[26] A. L. Boyer, "NAVAL RESPONSE TO A CHANGED SECURITY ENVIRONMENT: Maritime Security in the Mediterranean," *Source Nav. War Coll. Rev.*, vol. 60, no. 3, pp. 73–100, 2007, doi: 10.2307/26396850.

[27] "Naval Research and Development: A Framework for Accelerating to the Navy & Marine Corps After Next."

[28] S. L. Pettyjohn, "U.S. Global Defense Posture, 1783–2011," 2011.

[29] B. Work, "A New Global Posture for a New Era," 2014, Accessed: Sep. 29, 2020. [Online]. Available: https://www.defense.gov/Newsroom/Speeches/Speech/Article/605614/a-new-global-posture-for-a-new-era/.

[30] D. Mahadzir, "SECDEF Esper: U.S. Will 'Keep Up the Pace' of South China Sea Freedom of Navigation Operations ," 2020.

[31] Department of Defense, "National Defense Strategy," 2018. Accessed: Sep. 30, 2020. [Online]. Available:

https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.

[32] E. Freund, “Freedom of Navigation in the A Practical Guide,” *Harvard Kennedy Sch.*, pp. 1–41, 2017.

[33] S. Stashwick, “China Deploys Long-Range Anti-Ship and Anti-Air Missiles to Spratly Islands For First Time – The Diplomat,” *The Diplomat*, 2018.

[34] H. Beech, “China’s Sea Control Is a Done Deal, ‘Short of War With the U.S.’ - The New York Times,” 2018.

[35] K. P. Ekman, “Applying Cost Imposition Strategies against China,” *Strateg. Stud. Q.*, no. Spring, pp. 26–59, 2015, [Online]. Available: https://www.airuniversity.af.edu/Portals/10/SSQ/documents/Volume-09_Issue-1/ekman.pdf.

[36] R. Brodie, “The Navy Needs Coastal Defense Cruise Missiles,” 2017.

[37] B. T. Oder, “Security THE DIMENSIONS OF RUSSIAN SEA DENIAL IN THE BALTIC SEA,” 2018.

[38] A. Schmidt, “Countering Anti-Access / Area Denial,” vol. 49, no. 0, pp. 1–8, 2017.

[39] CCDCOE, “NATO Recognises Cyberspace as a ‘Domain of Operations’ at Warsaw Summit.” https://ccdcoe.org/incyder-articles/nato-recognises-cyberspace-as-a-domain-of-operations-at-warsaw-summit/ (accessed Oct. 09, 2020).

[40] R. O’Rourke, “Navy DDG-51 and DDG-1000 destroyer programs: Background and issues for Congress,” pp. 1–83, 2013.

[41] Joint Chiefs of Staff, “CHARTER OF THE JOINT REQUIREMENTS OVERSIGHT COUNCIL (JROC) AND IMPLEMENTATION OF THE JOINT CAPABILITIES INTEGRATION AND DEVELOPMENT SYSTEM (JCIDS) ,” 2018.

[42] Office of Aerospace Studies, “Capabilities-Based Assessment (CBA) Handbook,” vol. 5522, no. December, 2017.

[43] Joint Chiefs of Staff, “Deployment and Redeployment Operations,” no. January 2018, 2007.

[44] R. Edson and J. Frittman, “Illustrating the Concept of Operations (CONOPs) Continuum and Its Relationship to the Acquisition Lifecycle,” *Seventh Annu. Acquis. Res. Symp. Wednesday Sess. Vol. I*, pp. 238–257, 2010.

[45] P. K. Davis and Z. M. Khalilzad, “A composite approach to Air Force planning,” 1996.

[46] R. Neches, “Engineered Resilient Systems (ERS) S&T Priority Description and Roadmap,” 2011.

[47] M. Mazarr, K. Best, B. Laird, E. Larson, M. Linick, and D. Madden, *The U.S. Department of Defense’s Planning Process: Components and Challenges*. 2019.

[48] F. S. Larrabee, J. Gordon, and P. A. Wilson, “The Right Stuff: Defense Planning

Challenges for A New Century," *Source Natl. Interes.*, no. 77, pp. 50–58, 2004.

[49] C. Clausewitz, "On War," in *On War*, 1982, p. 142.

[50] J. Gideon, C. H. Dagli, A. K. Miller, J. M. Gideon, C. H. Dagli, and A. Miller, "Scholars ' Mine Taxonomy of Systems-of-Systems Taxonomy of Systems-of-Systems," 2005.

[51] J. Gideon, C. H. Dagli, A. K. Miller, J. M. Gideon, and A. Miller, "Taxonomy of Systems-of-Systems," *Proc. CSER*, 2005, Accessed: Oct. 26, 2021. [Online]. Available: https://scholarsmine.mst.edu/engman_syseng_facwork.

[52] J. Willison, "C4I Community of Interest C2 Roadmap Agenda," no. March, 2015.

[53] B. Martin *et al.*, "Advancing Autonomous Systems An Analysis of Current and Future Technology for Unmanned Maritime Vehicles," 2019.

[54] K. S. Brower *et al.*, "Comparative naval architecture analysis of nato and soviet frigates - part I," no. October 1980, pp. 87–99, 1980.

[55] L. Ferreiro and G. Autret, "A Comparison of French and U.S. Amphibious Ships," *Nav. Eng. J.*, vol. 107, no. 3, pp. 167–184, 1995, doi: 10.1111/j.1559-3584.1995.tb03044.x.

[56] Í. A. Fonseca and H. M. Gaspar, "Challenges when creating a cohesive digital twin ship: a data modelling perspective," *https://doi.org/10.1080/09377255.2020.1815140*, vol. 68, no. 2, pp. 70–83, 2020, doi: 10.1080/09377255.2020.1815140.

[57] Geographic information system, “Real World Map.” https://peiyunprocess.files.wordpress.com/2014/10/ca6ce94cdd2e09a1da8aa6ec22336835.jpg.

[58] R. Lempert, D. Warren, R. Henry, R. Button, J. Klenk, and K. Giglio, *Defense Resource Planning Under Uncertainty: An Application of Robust Decision Making to Munitions Mix Planning*. 2016.

[59] E. V Larson, “Force Planning Scenarios, 1945-2016: Their Origins and Use in Defense Strategic Planning,” *Force Plan. Scenar.*, 2019, Accessed: Oct. 27, 2021. [Online]. Available: www.rand.org/giving/contribute.

[60] Department of Defense, “2014 Quadrennial Defense Review,” 2014.

[61] J. F. Troxell, “Force Planning in an Era of Uncertainty: Two MRCs as a Force Sizing Framework,” *Strateg. Stud. Institute, U.S. Army War Coll.*, 1997.

[62] P. K. Davis and L. Finch, “Defense Planning for the Post-Cold War Era,” 1993.

[63] M. F. Fitzsimmons, “Scenario Planning and Strategy in the Pentagon.”

[64] R. J. Lempert *et al.*, “Making Good Decisions Without Predictions,” 2013, [Online]. Available: http://www.rand.org/pubs/research_briefs/RB9701.html.

[65] J. Wang, “A System Dynamics Simulation Model for a Four-rank Military Workforce,” 2007.

[66] R. R. Hill, J. O. Miller, and G. A. McIntyre, “Applications of discrete event

simulation modeling to military problems," *Winter Simul. Conf. Proc.*, vol. 1, pp. 780–788, 2001, doi: 10.1109/WSC.2001.977367.

[67] G. Arjan Veldhuis, N. De Reus, T. Logtens, G. Pallaske, and S. Carnohan, "The application of modelling and simulation in support of operational decision making during land operations."

[68] T. M. Cioppa, T. W. Lucas, and S. M. Sanchez, "Military applications of agent-based simulations," *Proc. - Winter Simul. Conf.*, vol. 1, pp. 171–180, 2004, doi: 10.1109/wsc.2004.1371314.

[69] C. J. Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, "Classification And Regression Trees (1st ed.). Routledge.," p. 368, 1984.

[70] W. Polonik and Z. Wang, "PRIM Analysis," 2007.

[71] A. Abu-Hanna, B. Nannings, D. Dongelmans, and A. Hasman, "PRIM versus CART in subgroup discovery: When patience is harmful," *J. Biomed. Inform.*, vol. 43, no. 5, pp. 701–708, Oct. 2010, doi: 10.1016/J.JBI.2010.05.009.

[72] J. H. Kwakkel and M. Jaxa-Rozen, "Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes," *Environ. Model. Softw.*, vol. 79, pp. 311–321, May 2016, doi: 10.1016/J.ENVSOFT.2015.11.020.

[73] J. H. Friedman and N. I. Fisher, "Bump hunting in high-dimensional data," *Stat. Comput.*, vol. 9, no. 2, pp. 123–143, 1999, doi: 10.1023/A:1008894516817.

[74] S. Dalal, B. Han, R. Lempert, A. Jaycocks, and A. Hackbarth, "Improving scenario

discovery using orthogonal rotations," *Environ. Model. Softw.*, vol. 48, pp. 49–64, Oct. 2013, doi: 10.1016/J.ENVSOFT.2013.05.013.

[75] A. M. Parker, S. V. Srinivasan, R. J. Lempert, and S. H. Berry, "Evaluating simulation-derived scenarios for effective decision support," *Technol. Forecast. Soc. Change*, vol. 91, pp. 64–77, Feb. 2015, doi: 10.1016/J.TECHFORE.2014.01.010.

[76] C. J. Van Rijsbergen, "INFORMATION RETRIEVAL."

[77] L. D. Ferreiro and M. H. Stonehouse, "A COMPARATIVE STUDY OF US AND UK FRIGATE DESIGN," *Br. Marit. Technol.*, 1991.

[78] L. Manfreda, "A systematic approach for military vessels crew sizing and its validation by comparison with in service naval ships," *NAV 2012 17th Int. Conf. Ships Shipp. Res.*, 2012, Accessed: Nov. 01, 2021. [Online]. Available: http://www.atenanazionale.it/nav/index.php/NAV/NAV2012/paper/view/64.

[79] C. J. Fitzgerald, N. R. Weston, Z. R. Putnam, and D. N. Mavris, "A Conceptual Design Environment for Technology Selection and Performance Optimization for Torpedoes," 2002.

[80] M. R. Kirby, D. N. Mavris, and S. Diego, "A Method for Technology Selection Based on Benefit, Available Schedule and Budget Resources," pp. 2000–2001, 2000.

[81] R. J. Lempert, "Robust Decision Making (RDM)," *Decis. Mak. under Deep*

Uncertain., pp. 23–51, 2019, doi: 10.1007/978-3-030-05252-2_2.

[82] C. Hamarat, J. H. Kwakkel, and E. Pruyt, "Adaptive Robust Design under deep uncertainty," *Technol. Forecast. Soc. Change*, vol. 80, no. 3, pp. 408–418, 2013, doi: 10.1016/j.techfore.2012.10.004.

[83] J. R. Kasprzyk, S. Nataraj, P. M. Reed, and R. J. Lempert, "Many objective robust decision making for complex environmental systems undergoing change," *Environ. Model. Softw.*, vol. 42, pp. 55–71, 2013, doi: 10.1016/j.envsoft.2012.12.007.

[84] R. J. Lempert, D. Warren, R. Henry, R. W. Button, J. Klenk, and K. Giglio, "Defense Resource Planning Under Uncertainty: An Application of Robust Decision Making to Munitions Mix Planning," 2016, Accessed: Nov. 01, 2021. [Online]. Available: www.rand.org/giving/contribute.

[85] V. Marchau, W. Walker, and P. Bloemen, "Decision making under deep uncertainty: from theory to practice," 2019, Accessed: Nov. 01, 2021. [Online]. Available: https://library.oapen.org/handle/20.500.12657/22900.

[86] M. Haasnoot, J. H. Kwakkel, W. E. Walker, and J. ter Maat, "Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world," *Glob. Environ. Chang.*, vol. 23, no. 2, pp. 485–498, Apr. 2013, doi: 10.1016/J.GLOENVCHA.2012.12.006.

[87] J. H. Kwakkel, W. E. Walker, and M. Haasnoot, "Coping with the Wickedness of Public Policy Problems: Approaches for Decision Making under Deep Uncertainty,"

J. Water Resour. Plan. Manag., vol. 142, no. 3, p. 01816001, Jan. 2016, doi: 10.1061/(ASCE)WR.1943-5452.0000626.

[88] R. Singh, P. M. Reed, and K. Keller, "Many-objective robust decision making for managing an ecosystem with a deeply uncertain threshold response," 2015, doi: 10.5751/ES-07687-200312.

[89] R. J. Lempert, D. G. Groves, S. W. Popper, and S. C. Bankes, "A general, analytic method for generating robust strategies and narrative scenarios," *Manage. Sci.*, vol. 52, no. 4, pp. 514–528, Apr. 2006, doi: 10.1287/mnsc.1050.0472.

[90] Q. Wang, M. Guidolin, and D. Savic, "Two-Objective Design of Benchmark Problems of a Water Distribution System via MOEAs: Towards the Best-Known Approximation of the True Pareto Front," 2015, Accessed: Nov. 24, 2021. [Online]. Available: http://hdl.handle.net/10871/16669.

[91] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, "A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1917, pp. 849–858, 2000, doi: 10.1007/3-540-45356-3_83.

[92] D. Hadka and P. Reed, "Borg: An Auto-Adaptive Many-Objective Evolutionary Computing FrameworkBorg: An Auto-Adaptive MOEA FrameworkD. Hadka and P. Reed," *Evol. Comput.*, vol. 21, no. 2, pp. 231–259, May 2013, doi: 10.1162/EVCO_A_00075.

[93] F. Zheng, A. R. Simpson, and A. C. Zecchin, "An efficient hybrid approach for multiobjective optimization of water distribution systems," *Water Resour. Res.*, vol. 50, no. 5, pp. 3650–3671, 2014, doi: 10.1002/2013WR014143.

[94] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining Convergence and Diversity in Evolutionary Multiobjective Optimization."

[95] G. Harik, E. Cantú, C. Cantú-Paz, D. E. Goldberg, and B. L. Miller, "The Gambler's Ruin Problem, Genetic Algorithms, and the Sizing of Populations."

[96] W. E. Walker, S. A. Rahman, and J. Cave, "Adaptive policies, policy analysis, and policy-making," *Eur. J. Oper. Res.*, vol. 128, no. 2, pp. 282–289, Jan. 2001, doi: 10.1016/S0377-2217(00)00071-0.

[97] W. J. Bradford, "The Theoretical Layered Air-Defence Capability of a Ship Engaged against Multiple Anti-Ship Capable Missile Attacks." Jun. 01, 1992, Accessed: Nov. 14, 2021. [Online]. Available: https://apps.dtic.mil/sti/citations/ADA258388.

[98] C. Laura Strickland, M. A. Day, K. DeMarco, L. Strickland, E. Squires, and C. Pippin, "Responding to unmanned aerial swarm saturation attacks with autonomous counter-swarms," *https://doi.org/10.1117/12.2305086*, vol. 10635, pp. 247–263, May 2018, doi: 10.1117/12.2305086.

[99] M. Gunzinger and B. Clark, "Winning the Salvo Competition."

[100] J. Benford, J. A. Swegle, and E. Schamiloglu, "High Power Microwaves," *High*

Power Microwaves, vol. 5776, no. 505, 2015, doi: 10.1201/b19681.

[101] R. O'Rourke, "Navy lasers, railgun, and hypervelocity projectile: Background and issues for congress," *Weapons Syst. Backgr. Issues Congr.*, pp. 47–105, 2018.

[102] S. Cid, "We don't need the S-band radar," *Insid. Navy*, vol. 23, no. 29, p. 10, 2010.

[103] L. Thomas and C. Moore, "CHIPS Articles: TALON — Robust Tactical Optical Communications," 2014. https://www.doncio.navy.mil/CHIPS/ArticleDetails.aspx?ID=5550 (accessed Oct. 12, 2020).

[104] A. Muller, "The Future of Naval Communications," 2010. https://www.naval-technology.com/features/feature87881/ (accessed Oct. 12, 2020).

[105] J. Schank, S. Savitz, K. Munson, B. Perkinson, J. McGee, and J. Sollinger, *Designing Adaptable Ships: Modularity and Flexibility in Future Ship Designs*, vol. D. 2016.

[106] D. B. Larter, "US Navy prepares major surge of littoral combat ship deployments," 2020. https://www.defensenews.com/naval/2020/07/31/the-us-navy-is-preparing-a-major-surge-of-lcs-deployments/ (accessed Oct. 12, 2020).

[107] M. Eckstein, "Navy Issues Final RFP for FFG(X) Next-Generation Frigate - USNI News," 2019. https://news.usni.org/2019/06/20/navy-issues-final-rfp-for-ffgx-next-generation-frigate (accessed Oct. 12, 2020).

[108] D. B. Larter, "If the US Navy isn't careful, its new unmanned tanker drone could

face a 3-year delay," 2020. https://www.defensenews.com/naval/2020/06/10/if-the-us-navy-isnt-careful-its-new-unmanned-tanker-drone-could-face-a-3-year-delay/ (accessed Oct. 13, 2020).

[109] W. Carroll and B. Hamblet, "Interview with the 'Air Boss' | Proceedings - September 2017 Vol. 143/9/1,375," vol. 143, no. 9, Sep. 2017, Accessed: Oct. 13, 2020. [Online]. Available: https://www.usni.org/magazines/proceedings/2017/september/interview-air-boss.

[110] J. Salak, "Weaponized Hovercraft for Distributed Lethality," 2015.

[111] M. McCulloch, "Roles for Up-gunned LCACs in Adaptive Force Packages," 2016.

[112] D. B. Larter, "With DoD's fleet of 2045, the US military's chief signals he's all-in on sea power," 2020. https://www.defensenews.com/naval/2020/10/06/with-its-fleet-of-2045-the-us-militarys-chief-signals-hes-all-in-on-sea-power/ (accessed Oct. 11, 2020).

[113] S. J. Freedberg, "DSD Work Embraces DARPA's Robot Boat, Sea Hunter « Breaking Defense - Defense industry news, analysis and commentary," 2016.

[114] D. B. Larter, "With billions planned in funding, the US Navy charts its unmanned future," 2019.

[115] L. Goldstein and S. Knight, "Wired for Sound in the 'Near Seas' | Proceedings - April 2014 Vol. 140/4/1,334," 2014, Accessed: Oct. 13, 2020. [Online]. Available: https://www.usni.org/magazines/proceedings/2014/april/wired-sound-near-seas.

[116] M. Peck, "Russia Wants To Use AI-Sea Mines To Sink America's Navy | The National Interest," 2020.

[117] R. Uppal, "Rising threat of Mine Warfare with development of stealthy, smart, maneuverable and networked sea mines | International Defense Security & Technology Inc.," 2019. https://idstch.com/threats/threat-mine-warfare-increasing-development-stealthy-smart-maneuverable-networked-sea-mines/ (accessed Oct. 13, 2020).

[118] D. B. Larter, "Here's what you need to know about the US Navy's new deadly (and expensive) attack subs," 2019.

[119] "French and British Submarines Collide - The New York Times," 2009.

[120] M. Eckstein, "Sonar Equipped Drone Fleets Could be Key to Future Submarine Warfare - USNI News," 2020. https://news.usni.org/2020/03/09/sonar-equipped-drone-fleets-could-be-key-to-future-submarine-warfare (accessed Oct. 14, 2020).

[121] Naval Postgraduate School, "Undersea Communications Between Submarines and Unmanned Undersea Vehicles in a Command and Control Denied Environment - Aerospace & Defense Technology," 2016. https://www.aerodefensetech.com/component/content/article/adt/tech-briefs/machinery-and-automation/24599 (accessed Oct. 13, 2020).

[122] S. Roblin, "Meet the 1 Submarine That Terrifies the U.S. Navy More Than Any Other | The National Interest," 2018. https://nationalinterest.org/blog/buzz/meet-1-

submarine-terrifies-us-navy-more-any-other-37922 (accessed Oct. 14, 2020).

[123] D. Majumdar, “US Navy’s Next Submarine: Super Stealthy and Now Underwater Aircraft Carrier? | The National Interest,” 2016.

[124] Office of Naval Research, “Large Displacement Unmanned Underwater Vehicle Innovative Naval Prototype (LDUUV INP) Energy Section Technology,” *Program*, vol. 102, no. July, pp. 1–26, 2011.

[125] M. Ball, “BAE Systems Integrates New Capabilities into Riptide UUVs | Unmanned Systems Technology,” 2020. https://www.unmannedsystemstechnology.com/2020/05/bae-systems-integrates-new-capabilities-into-riptide-uuvs/ (accessed Oct. 14, 2020).