2020-09

AI-AUGMENTED DECISION SUPPORT SYSTEMS: APPLICATION IN MARITIME DECISION MAKING UNDER CONDITIONS OF METOC UNCERTAINTY

Uziel, Steven J.

Monterey, CA; Naval Postgraduate School

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

AI-AUGMENTED DECISION SUPPORT SYSTEMS: APPLICATION IN MARITIME DECISION MAKING UNDER CONDITIONS OF METOC UNCERTAINTY

by

Steven J. Uziel

September 2020

Thesis Advisor: Mollie R. McGuire
Second Reader: Steven J. Iatrou

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2020	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE
AI-AUGMENTED DECISION SUPPORT SYSTEMS: APPLICATION IN MARITIME DECISION MAKING UNDER CONDITIONS OF METOC UNCERTAINTY

5. FUNDING NUMBERS

6. AUTHOR(S) Steven J. Uziel

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research, Arlington, VA 22203

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release. Distribution is unlimited.

12b. DISTRIBUTION CODE
A

13. ABSTRACT (maximum 200 words)

The ability for a human to overlay information from disparate sensor systems or remote databases into a common operational picture can enhance rapid decision making and implementation in a complex environment. This thesis focuses on operational uncertainty as a function of meteorological and oceanographic (METOC) effects on maritime route planning. Using an existing decision support system (DSS) with artificial intelligence (AI) algorithms developed by New Jersey Institute of Technology and University of Connecticut, cognitive load and time to decision were assessed for users of an AI-augmented DSS, accounting for METOC conditions and their effects, and users of a baseline, "as is," DSS system. Scenario uncertainty for the user was presented in the relative number of Pareto-optimal routes from two locations. Key results were (a) users of an AI-augmented DSS with a simplified interface completed assigned tasks in significantly less time than users of an information-dense, complex-interface AI-augmented DSS; (b) users of simplified, AI-augmented DSS arrived at decisions with lower cognitive load than baseline DSS and complex-interface AI-augmented DSS users; and (c) users relied mainly on quantitative data presented in tabular form to make route decisions. The differences found in user performance and cognitive load between levels of AI augmentation and interface complexity serve as a starting point for further exploration into maximizing the potential of human-machine teaming.

14. SUBJECT TERMS
meteorological and oceanographic, METOC, decision support system, DSS, artificial intelligence, AI, TMPLAR, cognitive load

15. NUMBER OF PAGES
127

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

AI-AUGMENTED DECISION SUPPORT SYSTEMS: APPLICATION IN MARITIME DECISION MAKING UNDER CONDITIONS OF METOC UNCERTAINTY

Steven J. Uziel
Major, United States Marine Corps
BS, Allegheny College, 2008

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
September 2020

Approved by: Mollie R. McGuire
Advisor

Steven J. Iatrou
Second Reader

Thomas J. Housel
Chair, Department of Information Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The ability for a human to overlay information from disparate sensor systems or remote databases into a common operational picture can enhance rapid decision making and implementation in a complex environment. This thesis focuses on operational uncertainty as a function of meteorological and oceanographic (METOC) effects on maritime route planning. Using an existing decision support system (DSS) with artificial intelligence (AI) algorithms developed by New Jersey Institute of Technology and University of Connecticut, cognitive load and time to decision were assessed for users of an AI-augmented DSS, accounting for METOC conditions and their effects, and users of a baseline, "as is," DSS system. Scenario uncertainty for the user was presented in the relative number of Pareto-optimal routes from two locations. Key results were (a) users of an AI-augmented DSS with a simplified interface completed assigned tasks in significantly less time than users of an information-dense, complex-interface AI-augmented DSS; (b) users of simplified, AI-augmented DSS arrived at decisions with lower cognitive load than baseline DSS and complex-interface AI-augmented DSS users; and (c) users relied mainly on quantitative data presented in tabular form to make route decisions. The differences found in user performance and cognitive load between levels of AI augmentation and interface complexity serve as a starting point for further exploration into maximizing the potential of human-machine teaming.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
ANOVA	Analysis of Variance
C2	command and control
CNO	Chief of Naval Operations
COA	course of action
COP	common operational picture
COAST	Courses of Action Simulation Tool
CONFIDENT	Conflict Identification
COP	common operational picture
DECISION MAKER	decision-maker
DOD	Department of Defense
DON	Department of the Navy
DSS	decision support system
GUI	graphical user interface
IS	information systems
METOC	meteorological and oceanographic
NJIT	New Jersey Institute of Technology
NRL–MRY	Naval Research Laboratory—Monterey
SA	situational awareness
TMPLAR	Tool for Multi-objective Planning and Asset Routing
UCONN	University of Connecticut
USN	United States Navy

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

The U.S. Navy (USN) plays a pivotal role in ensuring territorial integrity of treaty partners, supporting freedom of navigation for maritime trade, and maintaining global stability. As a key contributor to our nation's ability to project military power, the USN and specifically the Surface Warfare community often face dynamic and uncertain operational environments as it enforces U.S. policy in the vast and contested maritime warfighting domain. The breakdown of the post-Cold War world order and the reemergence of long-term, antagonistic competition by near-peer/peer rival states has the potential to degrade the primacy of U.S. influence in strategically critical operational theaters. Aside from the kinetic threats that will accompany a potential high-end conflict, factors including commercial traffic volume of confined waterways, current and projected meteorological and oceanographic (METOC) conditions, and the uncertainty that accompanies the fog of war add compounding layers of complexity in the planning and execution of modern surface actions. As testament to the complexity of maritime operations, the recent collision incidents involving both the USS *Fitzgerald* (DDG 62) and the USS *John S. McCain* (DDG 56) represent tragic but avoidable consequences of losses of situational awareness (SA) and failures to properly use available navigation tools (USN, 2017). In the case of the USS *Fitzgerald* and the USS *John S. McCain* collisions, the Office of the Chief of Naval Operations (CNO) found that ineffective command and control (C2), among other factors, contributed to the mishaps resulting in the loss of 17 U.S. Sailors (United States Navy [USN], 2017).

The USN, in efforts to both gain and maintain maritime superiority and minimize operational losses such as those described above, invests substantial time and financial resources in the training and education of Surface Warfare Officers (SWO) and enlisted leadership. However, as the operational environment evolves and becomes inherently more complex, the decision and execution cycle of naval leaders must further decrease the time required to arrive at appropriate, justified courses of action (COA). To this end, agile and effective planning in a dynamic environment requires the "right information from right

[sic] sources in the right context to the right person in the right time for the right purpose" (Smirnov, 2006, p. 30). Due to the myriad of external factors that impact not only the conduct of large-scale naval campaigns but also routine actions such as surface transits, digital Decision Support Systems (DSS) including Tool for Multi-objective Planning and Asset Routing (TMPLAR), Courses of Action Simulation Tool (COAST), and Conflict Identification (CONFIDENT) assist naval decision makers with data analysis, thereby streamlining the decision-making process. Recently, artificial intelligence (AI) algorithms have been proposed as a means to condense the decision-making cycle through interfacing with currently in use DSS systems. In concert with the University of Connecticut (UCONN), New Jersey Institute of Technology (NJIT), and Naval Research Laboratory—Monterey (NRL–MRY), this study evaluated the effectiveness of human decision-making aided by an AI-enhanced DSS in simulated surface transit scenarios.

B. PROBLEM STATEMENT

In the projected operational environment, the context-independent DSS that are currently employed provide an incomplete understanding of available sensor data and potentially limit the range of operational COAs available to U.S. planners in the maritime domain. Extant planning tools and current and C2 systems do not possess the ability to forecast spatial/temporal conflicts in surface routes, compile and process multi-source collections or sensor inputs in real time, or integrate the resultant data into a common operational picture (COP). This lack of multi-domain awareness resulting from a fragmented and flawed understanding of the battlespace and its effects curtails decision makers' ability to effectively mitigate risk, seize the tactical or operational initiative, and employ the force in the most resource efficient and tactically effective manner. As a result, naval planners currently lack the ability to rapidly develop, validate, and implement discrete COAs that account for all data that is theoretically available to the decision maker. The employment of AI to fuse, integrate, and display available sensor data utilizing current DSS applications' graphical user interfaces (GUI) and display architecture will allow the decision maker to account for fluid temporal, environmental, and adversarial considerations.

C. PURPOSE STATEMENT

The purpose of this quantitative study is to evaluate the effectiveness of UCONN/NJIT-developed, AI-augmentation to the TMPLAR application for surface vessel route planning/deconfliction and to assess its ability to reduce the cognitive workload of system users. The DSS framework currently under development aims to provide the user with real-time knowledge derived from environmental context and situational variables, potentially reducing the analytic burden for the user and facilitating the development and selection of situationally appropriate and tactically sound COAs. This experiment is important because, if it is successful, the employment of an AI-augmented, proactive DSS could reduce costs associated with greater fuel expenditures due to inefficient route planning while underway, mitigate contributing factors to safety mishaps at sea, and allow individual ships to spend more time influencing the battlespace vice transiting to/between theaters. From the use of environmental circumstances and surface transits as the parameters for this initial augmented DSS application and study, the framework could later be extended to consider unknown/enemy threats and additional domains (subsurface, air, etc.), or applied to land-based campaigns in expeditionary or mature operational theaters.

D. RESEARCH QUESTIONS

1. How does the employment of a context-aware, AI-augmented DSS impact decision-making timelines over baseline DSS in simulated ship transits under conditions of METOC uncertainty?

2. How can the employment of a context-aware, AI-augmented DSS affect the cognitive load of a user over baseline DSS during surface transit planning in conditions of METOC uncertainty?

3. What information do AI-augmented and baseline DSS users prioritize when selecting Pareto-optimal routes during surface transit planning in conditions of METOC uncertainty?

E. RESEARCH METHOD

Utilizing context-dependent AI algorithms developed by NJIT and UCONN, ship navigation and routing performance decisions both with and without the use of an augmented DSS framework were evaluated and analyzed at NPS (Glasgow Hall, Room 103). The context-dependent, algorithm-augmented DSS was tested for efficacy by human decision makers under uncertain METOC conditions in a human-in-the-loop experiment. The experiment was structured as a 3 x 3 repeated measures design. The first independent variable was the employment of an AI-algorithm enhanced DSS or the baseline, "as is," DSS system (TMPLAR). Each participant was randomly assigned the use of an AI-enhanced or baseline, non-augmented DSS to accomplish the assigned tasks. The second independent variable was uncertainty of METOC conditions and was defined by the number of Pareto-optimal routes from two known locations given by the system. Levels of uncertainty were categorized as: 1) Low/Easy difficulty: decision maker was given choice of COAs with projected fuel usage within one standard deviation of the mean for the given route, 2) Moderate/Medium difficulty: decision maker was given choice of COAs with projected fuel usage within two standard deviations of the mean for the given route, 3) High/Hard difficulty: decision maker was given choice of all COAs across the entire distribution of fuel usage for the given route. Under each condition of METOC uncertainty (low, moderate, high), participants planned 20 routes using the TMPLAR application, totaling 60 routes per participant. Surface routes planned for the experiment were Alaska-Seattle, Norfolk-Haiti, Alaska-San Diego, and Gibraltar-Norfolk. Results were then compared between AI-augmented DSS and baseline DSS trials across the relative uncertainty conditions.

Pilot testing of AI-algorithm augmented and baseline TMPLAR application was conducted prior to experimentation in order to determine time required for participant familiarization and gauge to scenario difficulty. Results from the pilot test were used to shape expectations for the timeline of individual participant trials. After pilot testing was completed, the study's experiment was conduct with revised timelines gleaned from the pilot test. Experiment participants were drawn from the NPS student body with no

preference given to Department of Defense (DOD) service or warfare specialty. Participant trials were initially projected to last 75 minutes and not exceed two hours in duration. Upon arrival in Glasgow 103, participants were given a verbal orientation on the experimental goals and design. At that time, the participants provided written informed consent to participate in the experiment. Participants were then seated in front of a monitor equipped with an eye tracking and pupillometry measurement system. Participants then spent approximately one minute allowing the system to calibrate to their eye movements. This was then followed by an orientation of the TMPLAR user interface and hands-on practical application of the requisite application functions. After the participants were comfortable with the task and the expectations of the experiment, 60 surface routes were selected across three degrees of METOC uncertainty which were presented to them on a semi-random basis. Simultaneous to the conduct of route selection, eye tracking and pupillometry data was collected from the participants to gauge cognitive workload between users of augmented and baseline DSS applications. Upon completion of the route planning, participants populated questionnaires relating to demographics and their experiences while conducting the experiment.

At the conclusion of the experiment, the data collected were analyzed to determine the relationship between route optimization and use of AI-augmented DSS under conditions of METOC uncertainty and the cognitive workload associated with each. Recommendations and conclusions proposed and further areas of research into the subject are recorded in Chapter V of this thesis.

F. DATA, OBSERVATIONS, AND ANALYSIS METHOD

The data presented to the user were structured corresponding to distance traveled, time of passage, and total fuel consumption between two geospatial data points per trial. Experimental observations were collected by the recording of user-selected, DSS-generated routes simulated within TMPLAR. The number of iterations of the TMPLAR simulation addressing each independent variable (uncertainty level and employment of AI-augmented/baseline DSS systems) were regulated by the researcher to ensure that the sample is representative of the target user population.

A 3 (DSS) x 3 (uncertainty) Analysis of Variance (ANOVA) was utilized to compare sample means and covariance across the dependent variables (time, optimal route, number of routes considered). This procedure was used to determine if changes in the independent variables had significant effects on the dependent variables and if there were any causal relationships among the dependent or independent variables. Further analysis was done with the eye tracking and pupillometry data. Eye-tracking data can provide insight into the decision-making process (e.g., what information was considered), while pupillometry is commonly used as a measure indicating cognitive load (e.g., Coyne et al., 2009; Klinger et al., 2010).

G. POTENTIAL BENEFITS, LIMITATIONS, AND RECOMMENDATIONS

The benefits of this study to the USN is twofold. First, AI-augmentation to existing DSS systems exhibits the potential to improve the operational mobility of surface vessels through more efficient routing, more complete exploitation of existing environmental data, and cost savings through minimized fuel consumption. Additionally, the USN will benefit from the ability to condense the planning and decision cycle of key decision makers through the fusion, analysis, and display of disparate sensor data and the reduction of the human cognitive workload, thereby enabling leaders to make an informed decision with the best information available faster than the adversary.

Due to the parameters of the AI algorithms developed by NJIT and UCONN, only five origin/destination pairs were trialed during the experiment. These routes link routinely traveled points of origin and destinations with little intervening landmasses or terrain, meaning it is possible to follow largely direct courses for large portions of the simulated voyages. Recommendations for future research will include experimentation testing the efficacy of AI-augmented DSS systems in more complicated and dynamic scenarios.

H. THESIS ORGANIZATION

This thesis is organized into five chapters. Chapter I outlines the background, problem, and purpose of the study. Chapter II contains a literature review of the current body of knowledge relating to METOC effects on maritime operations, naval C2, AI, and

the TMPLAR application. Chapter III describes the design methodology and experiment procedure. Chapter IV communicates the results of the experiment, and Chapter V provides the conclusions and recommendations.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

A. INTRODUCTION

The body of literature addressing the fields of human decision-making and AI is large and robust. From the available information in existing studies, it is apparent that the employment of AI will play increasingly larger roles in how humans weigh alternatives and make comparative decisions. With the ever-expanding capabilities of computer hardware and processing power due to the effects of Moore's Law and the increasing sophistication and deployability of information systems (IS), real-world applications and benefits of AI employment are being pushed to the tactical edge to meet the needs of the warfighter. As such, this chapter will discuss how METOC factors affect maritime operations, naval C2 and DSS employment, AI, the notion of cognitive load, and the TMPLAR application.

B. METOC

1. Environmental Factors and Their Effects on Maritime Operations

METOC is the term used by the DOD to encompass, "all environmental factors, from the sub-bottom of the Earth's oceans through maritime, land areas, airspace, ionosphere, and outward into space" (Joint Chiefs of Staff [JCS], 2020, p. 140). In a military context, METOC effects have the capacity to impact all combatants, operations, and equipment employed in a given battlespace (United States Marine Corps [USMC], 2018b). As more advanced weapons and support systems are adopted by the force, knowledge of current and projected METOC conditions are increasingly critical to not only the function of these systems but in the conduct of missions and campaigns as a whole (USMC, 2018b). As a result of the Department of the Navy's (DON) mission set and its ability to operate in and influence the maritime (surface and subsurface), land, air, and space domains, an accurate understanding of the METOC environment by all mission stakeholders is critical to the enduring success of the naval services.

METOC factors, both positive and negative, have great effects on the ability of a force to conduct both routine movements of afloat assets and aggregated maritime

operations. The concept known as “weather routing” is a collective term used to describe the actions taken to plan and safely execute a voyage that accounts for various constraints, including environmental factors, ship-dependent goals, and cargo care (Krata & Szlapczynska, 2018). The goal of weather routing is to minimize the resistance of a ship as it transits through the waterspace, limit undesired motion disrupting personnel and cargo, and the seeking of calmer conditions where possible to optimize ship performance (Perera & Soares, 2017). Methods of weather routing typically account for both time-independent and static constraints such as landmasses, well-known hazards, and territorial waters and dynamic, time-dependent factors such as safety of passage thresholds for METOC conditions (Krata & Szlapczynska, 2018). As a result, the practice of weather routing is done to offset the hazards and negative effects of degraded environmental conditions while optimizing time of passage or fuel economy in a maritime movement.

The three most basic environmental factors that affect the movement of ships while underway are winds, waves, and currents (Cai et al., 2014). Wind, in the context of maritime movements, can have both positive and negative effects for the movement of a vessel (Cai et al., 2014). In cases of wind speeds of less than 20 knots, an underway vessel will typically gain speed if it is experiencing a tailwind or lose speed if moving into a headwind (National Geospatial-Intelligence Agency [NGA], 2017). By extension, it is also expected that a ship would either expend more or less fuel, accordingly, if a given speed is to be maintained as one navigates through a headwind or tailwind. At wind speeds of greater than 20 knots, ship speed is typically reduced regardless of wind direction due to increased wave activity, resulting in the need for greater frequency of steering corrections (NGA, 2017). Furthermore, the concept of “sail area” is of significance to underway ships dealing with winds. Sail area refers to how much of a ship’s surface area is exposed to the relative wind (NGA, 2017). Consequently, vessels with greater sail area as a ratio to overall length are more greatly affected by the force of head, tail, and beam winds (NGA, 2017).

According to the *American Practical Navigator, Volume I* published by the NGA, “wave height is the major factor affecting ship performance” (NGA, 2017, p. 695). Much like the effects of wind, wave actions moving in an opposite direction to the heading of a ship result in reduced speed, while those moving with the ship from the stern can increase

speed up to a certain point, beyond which performance is negatively affected (NGA, 2017). In general, large and sustained waves will result in slower speeds for a given power input due to reduced propeller thrust and increased drag as a result of steering actions (NGA, 2017). For the case of large vessels, the effects of waves and swell are generally larger than that of wind speed and direction (NGA, 2017). The variables of wave height, wave angle, and period all have negative impacts on the resistance of a ship as it transits the waterspace (Perera & Soares, 2017).

As opposed to wind and significant waves, currents are not as problematic in the routing of vessels under most conditions, especially when transiting near low latitudes (NGA, 2017). Due to worldwide oceanic currents being largely known, many vessels routinely take advantage of these more predictable conditions; for example: westward movements between the Panama Canal and southwest Asia can be optimized with favorable currents by travelling along a latitude of roughly 22°N, thereby offsetting the time penalty that would otherwise be accepted due to increased distance (NGA, 2017). Aside from periodic and generally forecasted weather anomalies such as ElNiño, hurricanes, and tropical typhoons, currents are generally more predictable than winds or waves (NGA, 2017).

2. U.S. Navy Approach to METOC Support

U.S. Joint Doctrine as it pertains to METOC is governed by four principles (JCS, 2018b). The first of these is accuracy, which attests that the data provided must be, "measurably correct in representing the current and future state of the environment," (JCS, 2018, p. vii). The next principle is consistency, which means that data provided to all echelons of the supported command must be synchronized with no conflicting information (JCS, 2018b). The third joint principle of METOC is relevancy, which is associated with the information's usefulness in the given situation (JCS, 2018b). The final joint METOC principle is timeliness, describing that information must be provided at a point early enough to where it can provide input into the commander's decision cycle (JCS, 2018b). Nested under the joint METOC principles are the METOC processes, which are defined as: collection, analysis, prediction, tailoring, dissemination, integration, and mitigation (JCS,

2018b). These processes in turn support the characterization and exploitation functions, in which characterization is the description of the past, present, or future METOC state and exploitation which represents the mitigation of negative effects and the capitalization on advantageous conditions (JCS, 2018b). Refer to Figure 1 for an illustration of the relationship between METOC processes and functions and their support to the decision cycle.

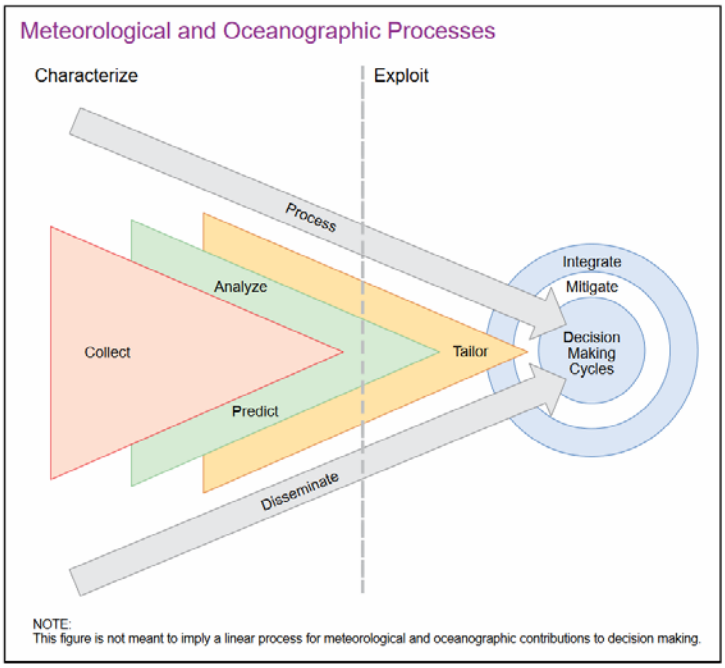


Figure 1. METOC Processes and Functions. Source: Joint Chiefs of Staff (2018b).

Established under the framework set by Joint Doctrine, the DON possesses considerable organic capabilities to characterize and exploit METOC information. The organization that provides these capabilities in the DON is Naval Meteorology and Oceanography Command (NAVMETOCCOM) (USN, 2011). As an Echelon III command, Commander, NAVMETOCCOM (COMNAVMETOCCOM) provides general METOC, Bathymetry/Hydrography (Bathy/Hydro), and Precise Time and Astrometry (PTA), through four directorates: Undersea Warfare, Expeditionary Warfare, Weather Services, and Positioning, Navigation, and Timing (USN, 2011). These directorates provide

tailored products in support of Naval Oceanography Operations Command (NAVOCEANOPSCOM), which in turn provides products to both the fleet and Joint Force (USN, 2011). Aiding the directorates with production and subject-matter expertise are three production centers: Fleet Numerical Meteorology and Oceanography Center (FLENUMMETOCCEN), Naval Oceanographic Office (NAVOCEANO), and the U. S. Naval Observatory (USNAVOBSY) (USN 2011). Additionally, the Naval Meteorology and Oceanography Professional Development Center serves as the training establishment for NAVMETOCCOM. Refer to Figure 2 for an illustration of the organization of NAVMETOCCOM.

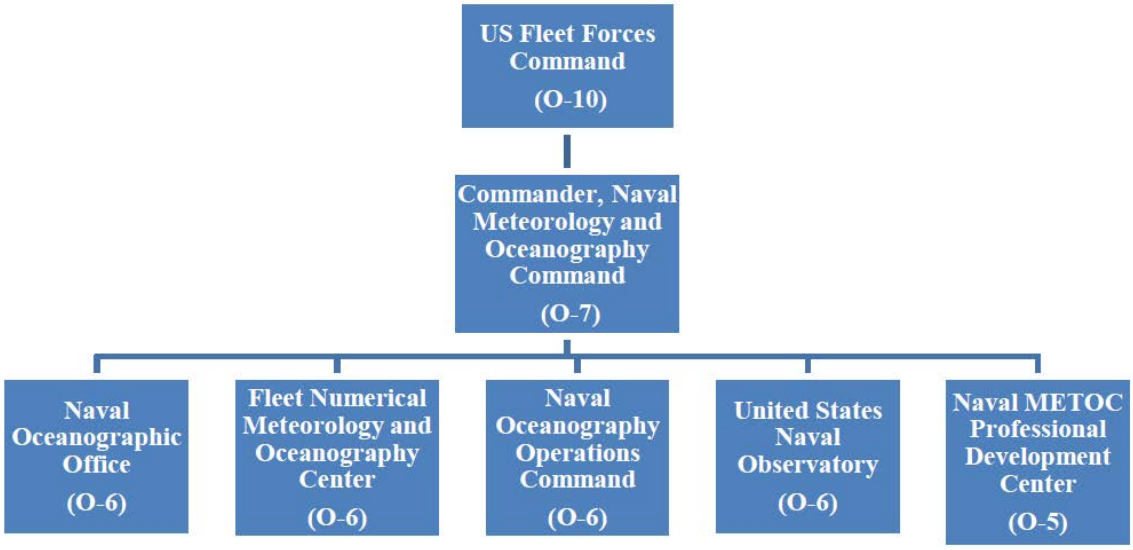


Figure 2. NAVMETOCOM Organizational Structure. Source: United States Navy (2011).

The products developed and services provided by the aforementioned organizations are made available to the force through a variety of mechanisms. Fleet Weather Center (FWC), Norfolk and FWC, San Diego, provide METOC maritime support to "enable fleet safety and readiness through accurate and timely weather forecasts, warnings and recommendations" (USN, 2011, p. 3–1)). The FWCs provide a myriad of products on request of a staff or ship, including Optimum Track Ship Routing (OTSR), Enroute Weather Forecasts (WEAX), High Winds and Seas Warnings, and several other products (USN, 2011). OSTR is

a message that may contain the following: weather advisories for conditions that approach or exceed safety thresholds for the supported unit, divert recommendations, and route recommendations given the type of ship, operational constraints, and environmental factors (USN, 2011). The WEAX is a weather and sea state forecast tailored to the requester's route or port area while the High Winds and Seas Warnings provide notification of areas exhibiting winds greater than or equal to 35 knots and seas greater than or equal to 12 feet (USN, 2011). These products are made available on a push and pull basis and contribute to the safe accomplishment of assigned tasks while underway.

Aside from the functions of the FWCs, METOC support is pushed to the operational forces through Fleet Operations Support. Fleet Operations Support provides "timely, comprehensive and tactically relevant METOC products and services in direct support of deploying Carrier Strike Group (CSG), Expeditionary Strike Group (ESG), and Amphibious Readiness Group (ARG) Commanders, assigned units, staff and other U.S. and Joint or Coalition forces, as directed" (USN, 2011, p. 5–1)). This is accomplished through the attachment of Strike Group Oceanography Teams (SGOT) to deploying CVN and LHA/D platforms (USN, 2011). The SGOT assists in the mission planning for the strike group and augments organic METOC division staffs. When a condition arises that a ship or unit without organic METOC capability deploys, a smaller Mobile Environmental Team (MET) may be assigned, according to mission priority, to provide METOC services (USN, 2011).

C. NAVAL C2

1. Doctrinal Basis

Like the other warfighting functions, C2 is firmly founded in doctrine and institutional experience. *Joint Publication (JP) 3–32* describes C2 as "the means by which a commander synchronizes and/or integrates joint force activities" (JCS, 2018, p. I-2)). In all operational contexts, C2 is performed through an interrelated system of people/organizations, processes and procedures, and systems that are employed to not only plan operations, but direct and control elements toward the accomplishment of a mission (JCS, 2018). In the view of the USN, the principal element of C2 is command, which refers to the legal authority a commander holds due to rank or billet and the responsibility to

effectively employ the resources under their charge in the accomplishment of an assigned mission (USN, 1995). On the other hand, control is the concept through which a commander monitors and influences actions toward the accomplishment of a mission (USN, 1995). Using feedback mechanisms, a commander can keep apprised of the current situation, adapt to the external trends and forces in the battlespace, apply resources where needed, and synchronize the force (USN, 1995). As a result, C2 is an interactive process in which the leaders and the led are complementary pieces of the whole and can better react to changing conditions and take advantage of fleeting opportunities (USMC, 2018a).

Throughout the history of armed conflict, and in some military forces in the world today, the idea of C2 as a reciprocal, interactive process has not found universal acceptance. The inverse of the DON's concept of a reciprocal C2 relationship is a unidirectional flow of C2 (USMC, 2018a). In such a scheme where both command and control flow from top down, the overall system is less adaptable to changes and disruptions and is less likely to capitalize on emerging opportunities. Refer to Figure 3 for a comparison of C2 views.

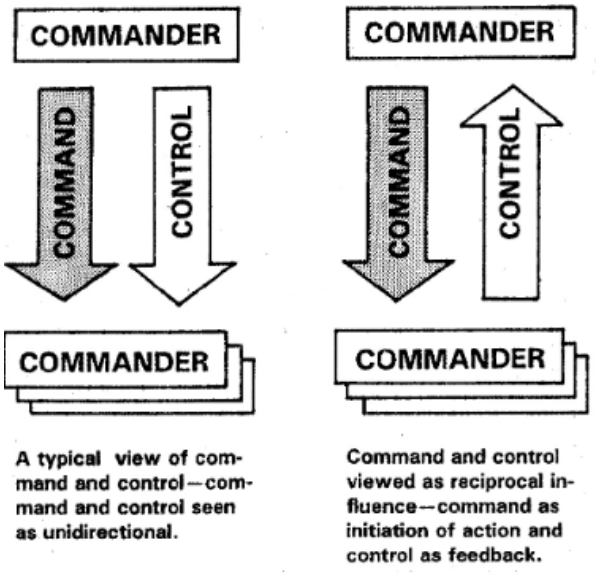


Figure 3. Views of the Relationship between Command and Control. Source: USMC (2018).

The vastness of the maritime domain, the complexity of the USN's mission, and its culture of independence require naval commanders to execute their orders with a high reliance on commander's intent and mission-type orders (JCS, 2018a). Critical to the successful implementation of mission-type orders is the commander's decision and execution cycle, commonly referred to as the Observe, Orient, Decide, Act (OODA) Loop (USN, 1995). Since developed by John Boyd in the 1980s, the model has become ubiquitous in the DOD in the 21st Century. In this model a decision maker first makes an observation of the environment through a variety of means including visual, sensor, combat report, etc. (USN, 1995). This data is then fused, overlaid, or otherwise displayed in some fashion that allows the decision maker to make a mental representation of the battlespace, thus orienting him or herself (USN, 1995). From the understanding gained from the mental picture, the decision maker then decides on a COA and issues the order or acts on the decision (USN, 1995). Refer to Figure 4 for an illustration of the decision and execution cycle in a military context.

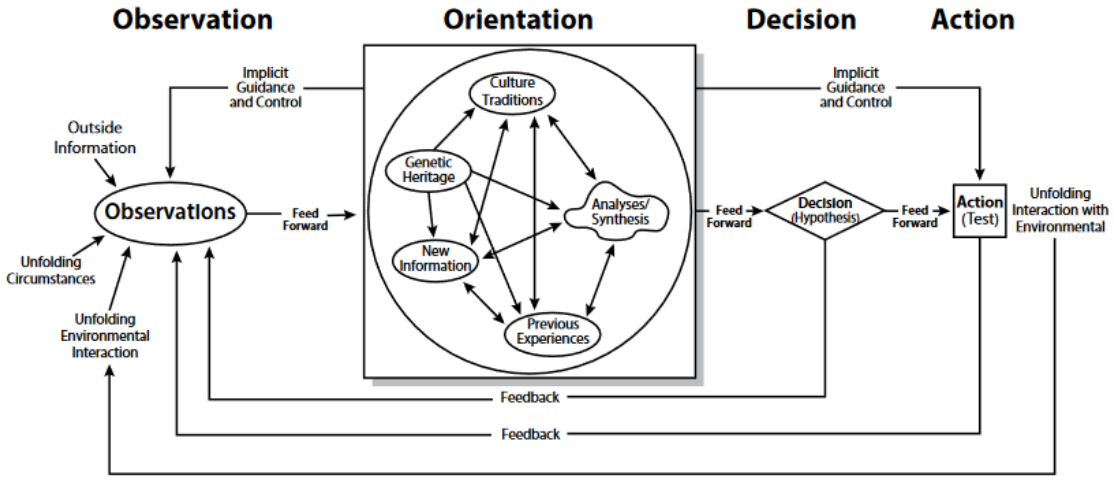


Figure 4. Decision and Execution Cycle. Source: Boyd (1986).

When considering the observe and orientation phases of the decision and execution cycle, one should take note of the subtleties associated with the terms data, information, knowledge, and understanding. These discrete concepts when arranged in a progressive, upward fashion make what is known as the Cognitive Hierarchy (USN, 1995). At the

lowest end of the progression, raw bits from sensors, signals, and visual observations comprise data, but are not necessarily meaningful in and of themselves (USN, 1995). Once it is processed, collated, or filtered, these data become information that begins to hold some limited value in a military context (USN, 1995). The evaluation, analysis, and fusion of information allows information to progress in the hierarchy to the state of knowledge, where the relationship between events can be derived (USN, 1995). Finally, understanding results from applied judgment and intuition, ultimately permitting synthesis and an awareness of a situation or system (USN, 1995). The ultimate goal of C2 is to achieve the state of understanding of an enemy's system or battlespace while allowing for a decision and execution cycle to operate inside that of the adversary. Refer to Figure 5 for a graphical depiction of the Cognitive Hierarchy.

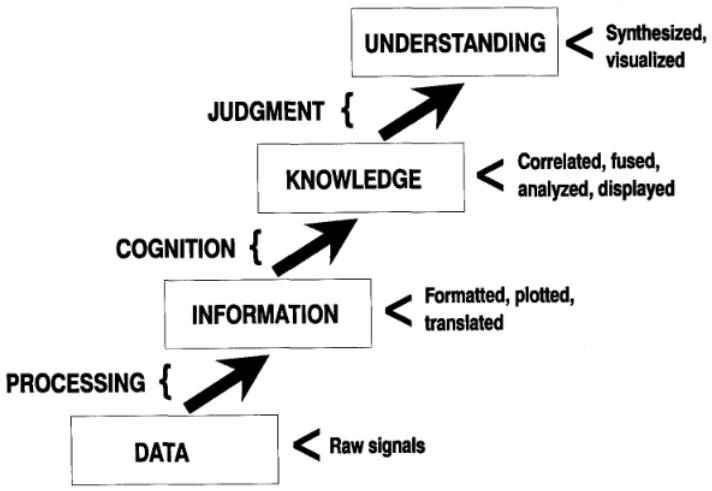


Figure 5. Cognitive Hierarchy. Source: United States Navy. (1995).

2. Decision-Making Theory

A popular theory of decision making that has gained traction in the U.S. military is Naturalistic Decision-Making (NDM). NDM arose in the 1980s as an attempt to conceptualize how people make decisions in real-world situations with emphasis on the role of the actor's experience (Klein, 2008). In the context of NDM study, the "real-world"

is an environment that is characterized by time constraints, ill-defined goals, and general uncertainty (Lipshitz et al., 2001). NDM holds that in real-world environments, people do not make decisions by comparing multiple options or by employing evaluation techniques, rather they use their own experiences to categorize a problem and formulate a COA (Klein, 2008). Lipshitz et al. describe the essential characteristics of NDM as, "proficient decision maker, situation-action matching decision rules, context-bound informal modeling, process orientation, and empirical-based prescription," (Lipshitz et al., 2001, p. 332). In the context of NDM, a proficient decision maker refers to an actor with the requisite experience or knowledge base in the area of the problem (Lipshitz et al., 2001). Situation-action matching decision rules describes the behavior of choosing an action because of a desirable rule-based outcome; for example, "Choose Option A because its outcome is better than Option B" (Lipshitz et al., 2001). Context-bound informal modeling describes the notion that abstract, formal models are of little utility because "expert knowledge," is domain- and context-specific (Lipshitz et al., 2001). Process orientation outlines the unique cognitive process of the decision maker and what information they actually use (Lipshitz et al., 2001). Finally, empirical-based prescription is the action of improving a decision maker's performance by "deriving prescriptions from descriptive models of expert performance" (Lipshitz et al., 2001, 335).

A well-known derivative of NDM is the Recognition-Primed Decision Making (RPDM). RPDM is unique in that it attests that patterns describe factors of causation for any given situation in the real world, thus are the basis for decision-making (Klein, 2008). These patterns recorded in the memory of the decision maker provide situational cues paired with expected outcomes that can be called upon when a real-world problem with similar circumstances is encountered (Klien, 2008). Much like NDM, expertise is stressed but the factors of analytical processing are also represented (Boyes & Potter, 2015). From pattern matching and past experiences, selections of COAs is often timely and spontaneous (Boyes & Potter, 2015). According to Klein, recognition, which drives pattern matching, has four aspects: plausible goals, expectancies, relevant cues, and possible actions that could be taken (Klein, 2008). Refer to Figure 6 for a visual model of RPDM.

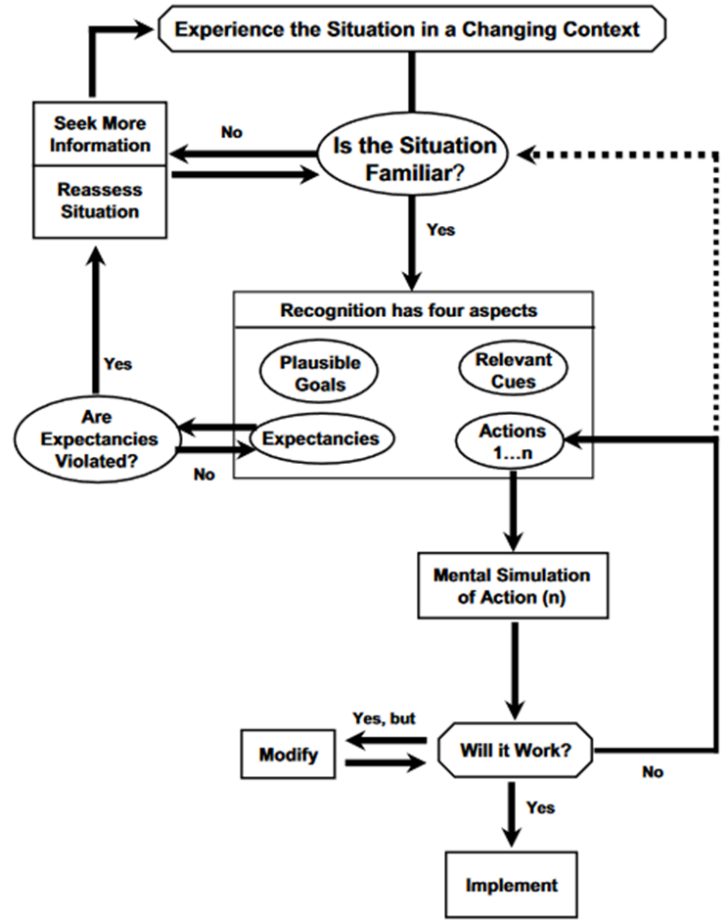


Figure 6. Recognition-Primed Decision Making. Source: Klein (1993).

Expanding from the basis of NDM, the Dynamic Model of Situated Cognition (DMSC) was developed to account for technology's inclusion in decision making processes in complex environments (Shattuck & Miller, 2005). DMSC arose to describe not only the processes decision makers use to arrive at decisions, but serves to illustrate the effects of technology on the perceptions in which decisions are based (Shattuck & Miller, 2005). Central to DMSC is the concept of situated cognition, which argues that knowledge and cognition are linked to the environment and context in which it was learned (Brown et al., 1989). Using this principle as a foundation, DMSC models the relationship between all data present in an environment, technological systems, perceptual and cognitive systems,

and the mental projection of the decision maker (Shattuck & Miller, 2005). Refer to Figure 7 for an illustration of the DMSC.

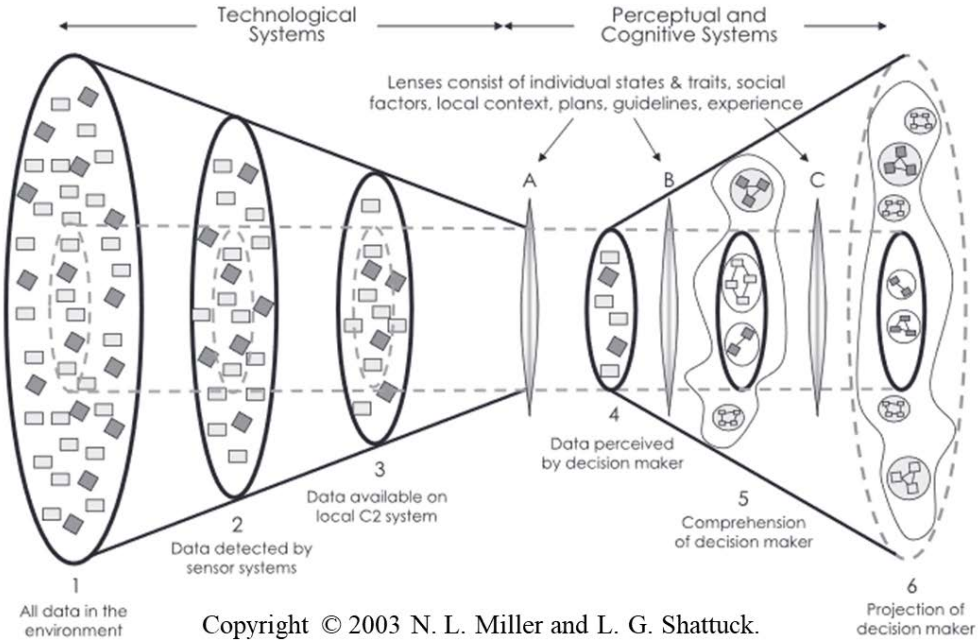


Figure 7. The Dynamic Model of Situated Cognition. Source: Shattuck and Miller (2005).

In Figure 7, the shapes in Oval 1 represent all of the information that exists in a given operational environment, hence a totally accurate but unobtainable view for a given moment in time (Shattuck & Miller, 2005). Progressing to the right in the diagram, Oval 2 represents the data that is detected or observed by a sensor or other technological system (Shattuck & Miller, 2005). The amount of data in Oval 2 will always be less than Oval 1 and will possibly contain inaccuracies, such as false positives or false negatives (Shattuck & Miller, 2005). Oval 3 contains the data in Oval 2 that is made available to the decision maker (Shattuck & Miller, 2005). Oval 3 will not contain all of the information present in Oval 2 and any errors manifested in Oval 2 will be represented in Oval 3 (Shattuck & Miller, 2005). At this point the data that has passed through the layers of the technological system is filtered through the first of three "lenses," that represent the factors of situated cognition and context before passing into the perceptual and cognitive systems (Shattuck

& Miller, 2005). Oval 4 represents the collection of data that is actually perceived by the decision maker, which then gets filtered through another lens before arriving at Oval 5, where it signifies comprehension of the data presented (Shattuck & Miller, 2005). Finally, the data passes through another lens and arrives at Oval 6 where the mental projection of the data is made by the decision maker (Shattuck & Miller, 2005). Like the other processes discussed, the DMSC is a continual process that occurs continuously as conditions change and new data is generated and made available.

3. DSS Systems / Multi-Objective Optimization

To deal with the high volume of data generated by today's combat systems, IS have been developed as a means to gather and process data in a way to provide value added to a military organization. As such, a DSS can be viewed as a subset of a the more general or generic IS classification. At a broad level, a DSS can be described as, "computerized systems to aid human decision makers by providing them better and more timely information, as well as the processing of this data in models," (Olson, 2013, para. 2). There exists multiple more specific, yet divergent definitions of the concept and as a result, several researchers have chosen to describe a DSS by its characteristics. In 1980, Sprague Jr. outlined the following as characteristics of a DSS:

- they tend to be aimed at the less well structured, underspecified problems that upper level managers typically face;
- they attempt to combine the use of models or analytic techniques with traditional data access and retrieval functions;
- they specifically focus on features which make them easy to use by noncomputer people in an interactive mode; and
- they emphasize flexibility and adapt- ability to accommodate changes in the environment and the decision making approach of the user. (Sprague, 1980, p. 2).

Stated otherwise, a DSS does not make a decision for a human decision maker, but allows for them to improve their mental processes and arrive at better and faster conclusions (Susnea, 2012).

Giving testament to the fact that the concept of a DSS is not a recent development, S.L. Alter described seven distinct DSS types, each of which exhibits distinct roles and

functions (Alter, 1976). The first of these are what are known as “file drawer systems,” which are computerized systems that provide immediate access to pre-organized information, such as in inventory or real-time monitoring of equipment. Secondly, “data analysis systems,” are platforms that allow non-managerial / non-technical personnel to conduct data retrieval, simple calculations, and general analysis without specialized training (Alter, 1976). The third DSS type, according to Alter, is “analysis information systems,” which exist to provide the user with the capability to access and analyze databases and small models, such as in sales information systems. “Accounting models,” the fourth DSS type described by Alter, have the purpose of forecasting the outcome of financial decisions based off of accounting principles. The fifth DSS, the “Representational model,” focuses on simulations based off of relationships that are not wholly definitional, such as in a risk analysis model (Alter, 1976). Alter’s sixth type, “optimization models,” provides the user with optimal solutions based off of known constraints. The final and seventh DSS type are what are known as “suggestion models,” which deal with structured tasks and are based both on rules and/or optimization (Alter, 1976). These seven DSS types are generally classified as data-oriented, containing the file drawer, data analysis, and analysis information systems, and model-oriented, including the accounting, representational, optimization, and suggestion models. Refer to Figure 8 for Alter’s general classification of DSS systems.

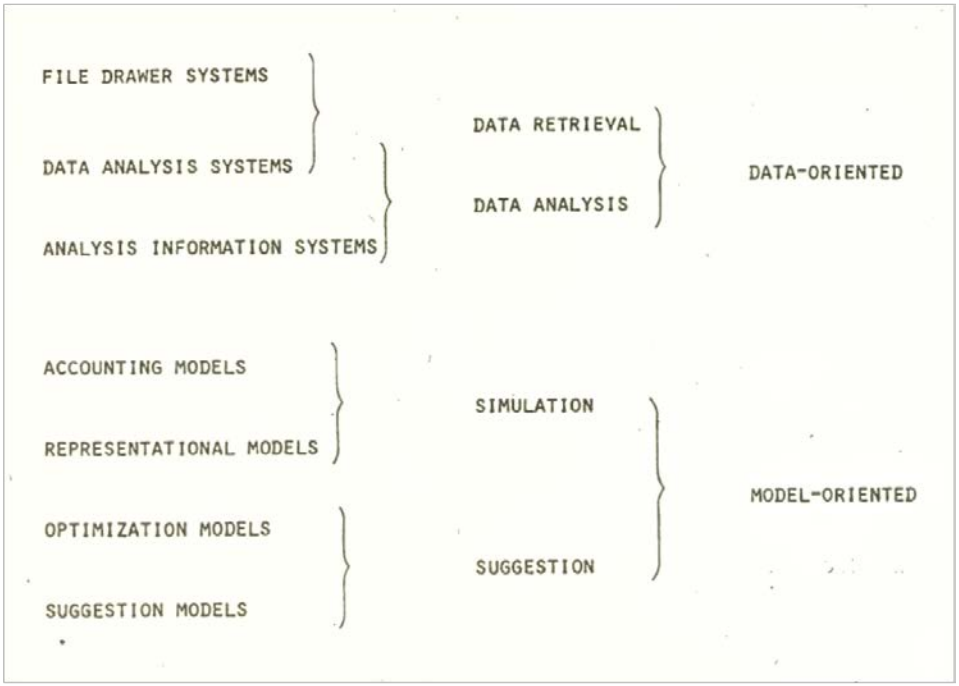


Figure 8. Data vs. Model System Types. Source: Alter (1976).

While all of these DSS types have military applications, the optimization model is of most concern to this study.

DSS tools used to fuse, structure, and present available data in a way conducive to rapid decision making do so by using multi-objective optimization. Multi-objective optimization is a discipline that involves decisions made that attempt to achieve two or more conflicting goals simultaneously (Ehrgott, 2008). The factors limiting how the decision maker achieves these goals are described as constraints, with those solutions not violating the constraints representing vectors of the feasible set (Ehrgott, 2008). The vectors of the feasible set in multi-objective optimization equate to the limits of the COAs available to the decision maker, either intuitively derived or as the result of DSS algorithms. Traditionally, two issues have the potential to impact the pertinence of optimization models: input data that is not exactly known and lack of a singular, clearly defined objective function (Ide & Schöbel, 2016). The process of military decision making, and in particular that in the maritime domain, closely resemble shortest path optimizations in which we seek to find an optimized solution to a problem with multiple antagonistic

objectives. In multi-objective shortest path optimizations, most often not all data are known (Ide & Schöbel, 2016). This gap in information is the result of lack of sensitivity of sensors, incorrect interpretation of available data, bandwidth restrictions, or general uncertainty of the battlespace, all cases which are commonplace in the current operational environment. The multi-objective method weighs the objectives proportionally according to subjective preferences and transposes the sum of the weighted vectors into a single objective function, which then produces multiple compromised, but optimized solutions (Chitra & Subbaraj, 2010). The application of AI to augment environmental data input, objective function calculation, and COA development and prioritization has the potential to enhance the fidelity of COAs produced, and produce more options as outputs from shortest path optimization algorithms.

A DSS that is in current widespread use in the USN is the Integrated Bridge System (IBS). IBS was developed by Sperry Marine, now part of Northrop Grumman, with the goal of maximizing bridge crew efficiency and safety (Sperry Marine, 2014). Within the USN alone, IBS has been fielded on CVN, CG, DDG, LHD, LPD, SSN, SSBN, LCAC, and PC platforms (Sperry Marine, 2014). Sperry Marine's IBS utilizes three core components to automate the collection, processing, and display of operational and environmental data: the Ship Control System (SCS), Automated Radar Plotting Aid (ARPA), and the Voyage Management System (VMS) (Sperry Marine, 2014). The basic functions of the SCS is to interface with Sperry Marine's Machinery Control System (MCS) and facilitate the monitoring and adjustment of a ship's propulsion system remotely (Sperry Marine, 2014). Sperry Marine's ARPA provides automation to the acquisition and tracking of contacts onto a common display (Sperry Marine, 2014). The VMS is IBS' navigation and movement planning system and provides functionality that fulfills Electronic Chart Display and Information System—Navy (ECDIS–N) requirements (Sperry Marine, 2014). The VMS component of IBS has several features that support decision-making, such as vessel advanced position prediction, a configurable conning information display, and a navigation station that allows for optional weather and route optimization software modules (Sperry Marine, 2010). Refer to Figure 9 for and illustration of a typical IBS configuration.

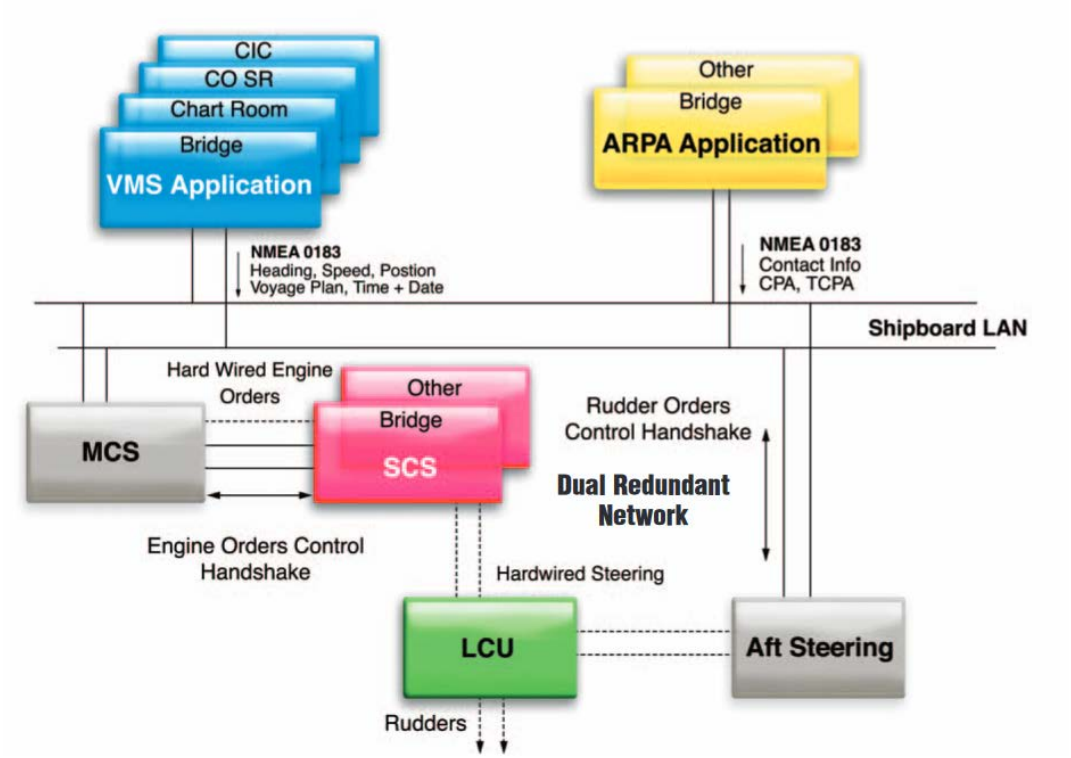


Figure 9. Military IBS Configuration. Source: Sperry Marine (2014).

D. ARTIFICIAL INTELLIGENCE

1. Relationship between AI and DSS

AI as a concept was first conceived in 1956 as "the ability of machines to understand, think, and learn in a similar way to human beings" (Pan, 2016, para. 2.1). The capabilities and possible applications of AI have grown since the concept's inception. Reflecting the current information environment consisting of the Internet, networked communities, integrated sensors, and big data, AI has introduced new computing architypes such as human-machine augmented intelligence and perception fusion (Pan, 2016). However, AI using computer-based tools to accomplish traditionally human-centric tasks does not equate to a system capable of decision making without virtualized emulation of how human beings arrive at decisions (Pomerol, 1997). Tasks are programmed and executed in accordance with instructions, reflecting preconceived decisions made by the system designer (Pomerol, 1997). In military functions, one of the most relevant applications of AI given today's operational environment is in DSSs. A DSS is a "look

ahead machine," that is a "multi-model, interactive system used by a decision maker to perform an exploration" (Pomerol, 1997, p. 21). As a result of system design and the assertion that rational decision makers can arrive at unique and independent solutions given the same problem, AI and by extension DSSs, encompass a large subjective component to their design and functionality (Pomerol, 1997). In the DSSs, the first phase that occurs within the system is diagnosis or pattern-matching, in which the perceived environment is compared and contrasted to recorded environments to gain an understanding of the current state (Pomerol, 1997). Once the system arrives at an understanding of the current state, the data are then weighed against other contributing factors producing a set of tenable outcomes, which are then subject to the system's preferences, producing a decision or chosen action (Pomerol, 1997). Additionally, human planners and DSSs, through different processes, share the ability to perform look-ahead reasoning or a "what if?" analysis in which one sees the consequences of a COA (Pomerol, 1997). A "what if?" analysis, otherwise known as scenario reasoning, has two essential outputs: 1) all outcomes possible given what is known of the situation or the environment, 2) the probability or plausibility of each outcome (Pomerol, 1997). The user may then apply his or her preferences based off of experience or personal/institutional biases and select a COA. Inherent to the notion of AI-augmentation to DSS is the concept of deep learning. Deep learning is a "technique for classifying patterns, based on sample data, using neural networks with multiple layers" (Marcus, 2018, p. 3). Utilizing input layers, hidden layers containing many nodes, and output layers, deep learning neural networks are often used as classification tools, in which the system decides what category given input parameters belong to (Marcus, 2018).

2. AI Applications of a Maritime DSS

Given the proliferation of AI and the benefits afforded to planners through the employment of DSSs, attention has recently been shifted to their applications in naval operations and maritime trade. The impacts of oceanographic effects on surface vessel routing is an inherent consideration to military and commercial maritime operations. Among the myriad of concerns of underway naval leaders are maintaining desired levels of fuel, speed, and the safety of ship and crew, all while considering external factors such planned fleet maneuvers and METOC effects (Chu et al., 2015). Also, due to the

complexities of the maritime domain, time of passage and total distance traveled represent additional objectives impacted by oceanographic phenomena (Sidoti et al., 2017). In gas turbine powerplants, fuel consumed increases approximately 1–4% at moderate speeds and roughly 9% at high speeds (Chu et al., 2015). Additionally, sea conditions have significant impacts on fuel consumption; at constant speed, a vessel will consume 10% more fuel when transiting in conditions of Sea State 4 and one knot surface currents than in calm conditions (Chu et al., 2015). In addition to surface hydrographic conditions, ship motions and wind resistance influence fuel economy and limit achievable speed (Vettor & Soares, 2016). As a result of naval movements being functions of multiple objectives (fuel efficiency, time available, speed, etc.), route optimization is inherently complex. Due to this complexity, DSS tools are required for human decision makers to optimize, evaluate, validate, and select COAs as they pertain to routing of surface vessels (Sidoti et al., 2017). With the vast amount of observable data that can potentially be integrated into a DSS from disparate sensors and sources, AI with the ability to learn has the capacity to further reduce the cognitive load of mission planners through the reduction or elimination of repetitive, time-consuming tasks.

E. COGNITIVE LOAD AND INDICATORS

1. Cognitive Load

AI has the capacity to reduce the cognitive load of mission planners through the reduction or elimination of repetitive, time-consuming tasks. Performance routine, repetitive tasks contributes to the "Intrinsic cognitive load," of users and places demands on their working memory. Intrinsic cognitive load is representative of the sum difficulty of a task and is compounded by the interactivity of its material or constituent factors (Coyne et al., 2009). The concept of intrinsic cognitive load is visible in the context of military operations where the complexity of the operational environment and the interrelationship of the actors within tax the working memory of planners as they attempt to conceptualize their surroundings as a prerequisite to developing COAs. As such, AI when applied to a DSS application can serve to minimize intrinsic cognitive load by automating routine or repetitive tasks, allowing the user to focus on making a sound decision with the best

information available in the shortest time while learning from the environment. To quantify this phenomenon, cognitive load of human users can be measured through the degree of pupil dilation while performing a task, also known as pupillometry (Klingner et al., (2010). In the past, cognitive load as indicated through pupillometry has shown to be lesser when users are presented with tasks given by visual rather than aural means, allowing individuals to omit the process of forming a mental picture of the task (Klinger et al., 2010). This suggests that better cognitive performance can be achieved through decreased workload on working memory, a state that context-dependent DSS applications aim to achieve for the user.

2. Eye Tracking / Pupillometry

As previously mentioned, the measure of pupil diameter has been used an indicator of cognitive load (Krejtz et al., 2018). Traditionally measured with specialized pupilometer equipment, modern eye tracking devices and software have emerged as an alternative method, allowing researchers to estimate cognitive load using changes to baseline pupil diameter in response to stimuli over time (Krejtz et al., 2018). Ikehara and Crosby attest that while pupil diameter is able to examine the range of cognitive states experienced by the user during the performance of a task, the results may be affected by presentation, such as sudden changes to the display intensity (Ikehara & Crosby, 2005). Aside from eye tracking, other indications of cognitive load that can be provided by eye trackers are gaze position, number and duration of fixations, and revisits, all of which indicate complexity or relative difficulty of the task (Ikehara & Crosby, 2005). In addition to pupil measurements and gaze data provided by eye trackers, the investigation of microsaccades, or involuntary eye movements during fixation, has been discussed as a measure of task difficulty perceived by a user (Krejtz et al., 2018). The magnitude of microsaccadic responses has been shown to increase with task difficulty along with intra- and inter-trial changes in pupil diameter (Krejtz et al., 2018). In addition, the eye movement and fixations recorded with an eye tracker have also shown to be a reliable indicator of task difficulty

3. User Manipulation of Data and Mouse Dynamics

Beyond recording of pupil and eye behavior, cognitive load has been measured by a concept known as mouse dynamics (MD), referring to the manner in which a user moves and uses a computer mouse during the conduct of a task (Grimes & Valacich, 2015). The basic user inputs that constitute a MD signature are screen coordinates of the mouse pointer, timestamps, or, "state changes," such as mouse movements or clicks (Grimes & Valacich, 2015). When a user interacts with a graphical environment, MD can be further categorized as general mouse movement, drag and drop, point and click, and silence indicating no movement (Ahmed & Traore, 2007). These types of positive clicking interactions with the user interface make logging user activity easier and more conducive to post-experimental data analysis. In 2015, Grimes and Valacich showed that when under high cognitive load users tend to move a mouse more slowly and with more changes in direction (Grimes & Valacich, 2015). In another study, Hibbeln et al. exploited the known connection to deception and increased cognitive activity and found that users committing simulated fraudulent activities in an experimental setting exhibited greater numbers of mouse clicks and higher measurements of other indicators of cognitive load (Hibbeln et al., 2014). Overall, past research shows that MD is a non-invasive, reliable indicator for cognitive load.

F. TMPLAR

1. Function / Capabilities

TMPLAR utilizes an automated system to suggest multiple route solutions to the human user, taking into account METOC information, geographic hazards to maritime traffic, bathymetry, and ship characteristics (Mishra et al., 2017). Recognizing that maritime movements involve multiple conflicting objectives such as time, fuel usage, and distance, TMPLAR presents multiple Pareto-optimal solutions to the user, allowing them to apply intuition and judgement as to the best compromise between the multiple objectives for a route solution (Mishra et al., 2017). TMPLAR was developed in the Python language and is imbedded natively with multiple empirical algorithms to achieve both single and multiple-objective optimization (Sidoti, 2018). As of August 2018, TMPLAR utilizes A*

and Q-Learning algorithms to accomplish single-objective optimizations and also Martins' and NAMOA algorithms for multi-objective optimizations (Sidoti, 2018). TMPLAR also, as of 2018, has four separate modes of operation: Lite, Robust, and Submarine which all employ a Fibonacci A algorithm and a Martins' mode that uses a Backwards Martins' algorithm (Sidoti, 2018). The four modes of operation in TMPLAR aggregate to support 18 current USN and Military Sealift Command (MSC) ship classes (Sidoti, 2018). Refer to Figure 10 for a complete list of supported ship classes.

1. CG-59
2. DD-963
3. DDG-58
4. DDG-90
5. DDG-93
6. JHSV-1
7. LCS-1
8. LCS-2
9. LHD-1
10. LSD-41
11. LSD-49
12. TAGS-60
13. TAKE-13
14. TAKR-300
15. TAKR-310
16. TAO-187
17. TAO-187 (Heavy)
18. TAOE-6

Submarine mode assumes a Los Angeles-class submarine (as of 8-14-2018).

Figure 10. Templar Supported Ship Classes. Source: Sidoti (2018).

2. Benefit of TMPLAR and Similar Systems

Context-dependent DSSs have been applied to real-world and virtual surface transits in the past. Recently, the TMPLAR application, used in concert with the embedded Smart Voyage Planning Decision Aid (SVPDA), was trialed against Ship Tracking and Routing Software++ (STARS++) in a shortest path problem set (Sidoti et al., 2017). Sidoti et al. found that through the addition of inclement weather avoidance algorithms to TMPLAR's functionality, the TMPLAR application outperformed software then in use by maritime navigators, resulting in a 33% fuel efficiency increase in simulated transits (Sidoti et al., 2017). In 2012, the SVPDA, with integrated data pertaining to wave height/period/velocity from WaveWatch III (WW3), ocean current velocity from U.S. Navy Coastal Ocean Model (NCOM), and surface winds from Navy Operational Global Atmospheric Prediction System (NOGAPS) was trailed in real-world surface passages

aboard two Arleigh Burke-class guided missile destroyers and one Henry J. Kaiser-class fleet replenishment oiler (Chu et al., 2015). This data, known as the METOC ensemble, showed the potential to achieve statistically significant fuel savings over great circle (GC) baselines when factors such as powerplant type and hull fouling are considered (Chu et al., 2015). Additionally, ocean current data obtained from the Surface Velocity Program (SVP) drifters over the period of 1985–2009 has shown to significantly reduce transit times for, "super-slow steaming," vessels in the western Pacific / east Asia region (Chang et al., 2013).

G. SUMMARY

The existing literature pertaining to DSS, AI, and oceanographic factors' impact on maritime operations provides a robust foundation for the evaluation of AI-augmented DSS and its applications in the maritime domain. The potential for budgetary savings due to decreased fuel expenditures, mitigated environmental impacts, and a more tactically and operationally agile force makes this research relevant to the USN's enduring mission set. Detailed analysis of the effectiveness and efficiency of AI-augmented DSS in uncertain environments will not only validate the concept for naval applications but provide insight into its applicability in other warfighting domains.

THIS PAGE INTENTIONALLY LEFT BLANK

III. METHODS

A. DESIGN AND PARTICIPANTS

1. Experimental Design

Assessing the efficacy of the context-dependent, AI-augmented DSS was accomplished utilizing a human-in-the-loop experiment. The experiment was structured as a 3 x 3 repeated measures design. The first independent variable in this construct the employment of AI-augmented or non-augmented DSS types. The first DSS type that was examined was classified as Level 0, indicating the as-is, off-the-shelf TMPLAR system with no AI-augmentation. Information that was made available to the user at Level 0 was restricted to a low-detail map depicting the start and end points of the voyage and a data table containing route and ship movement metrics. At Level 0, the user was not afforded any mechanism to conduct comparative analysis of ship routes other than the movement metrics displayed in the data table. Aside from the baseline system, the AI-augmented DSS was comprised of two additional DSS levels. Increasing in complexity of user interface from Level 0, Level 1 incorporated a user-customizable cumulative probabilistic distribution function, allowing for comparative analyses of fuel-usage or inclement weather probabilities based off of the underlying simulations. The visualization of the Pareto-optimal routes allowed for direct comparison of fuel usage and distance of travel of all routes calculated by the DSS. The Level 2 DSS level which represented an iterational model that was not adopted and therefore omitted from the experiment. The next level used in the experiment was the Level 3 DSS interface, and it was the most advanced, data-rich option. In addition to the baseline data present in Level 0 and the illustration of the cumulative probabilistic distribution function introduced at Level 1, the Level 3 interface incorporated a visualization of the Pareto-optimization of tenable sea routes as they were calculated using the AI algorithm embedded in the augmented DSS.

The second independent variable that was examined was the uncertainty of the simulated operational environment. Uncertainty for the purposes of this experiment was defined by the relative number of Pareto-optimal routes from the desired starting point and

destination per simulation. As discussed previously, levels of uncertainty were categorized as: 1) Low: decision maker given choice of COAs with projected fuel usage within one standard deviation of the mean for the given route, 2) Moderate: decision maker given choice of COAs with projected fuel usage within two standard deviations of the mean for the given route, 3) High: decision maker given choice of all COAs across the entire distribution of fuel usage for the given route. Participants were assigned a non-augmented (Level 0) or augmented (Level 1 or Level 3) in sequence of their participation resulting in a roughly equal number of participants across all DSS levels trialed. Within each trail, every participant was presented with 60 total navigation scenarios representing an approximately equal breakdown of Low, Moderate, and High uncertainty problems.

2. Participants

Participants for the experiment were drawn exclusively from the NPS student body. No particular service affiliation, prior experience, or participant nationality was required and all volunteers were accepted. The researchers used a variety of methods to recruit experimental subjects, all of which were approved by the NPS Institutional Review Board (IRB). Volunteers self-scheduled an experiment session through the www.signupgenius.com web application. Once a volunteer self-scheduled for a time slot, an email was automatically sent to the student researcher notifying him of the reservation.

In total 40 volunteers participated in the study. The service breakdown of experiment participants indicated a large majority of DON volunteers, generally reflecting the service breakdown of the NPS student body. Of the 40 volunteers, 21 had a USMC service affiliation and 13 held positions in the USN, totaling a combined 86% of all participants. Refer to Figure 11 for a complete experiment participant service breakdown.

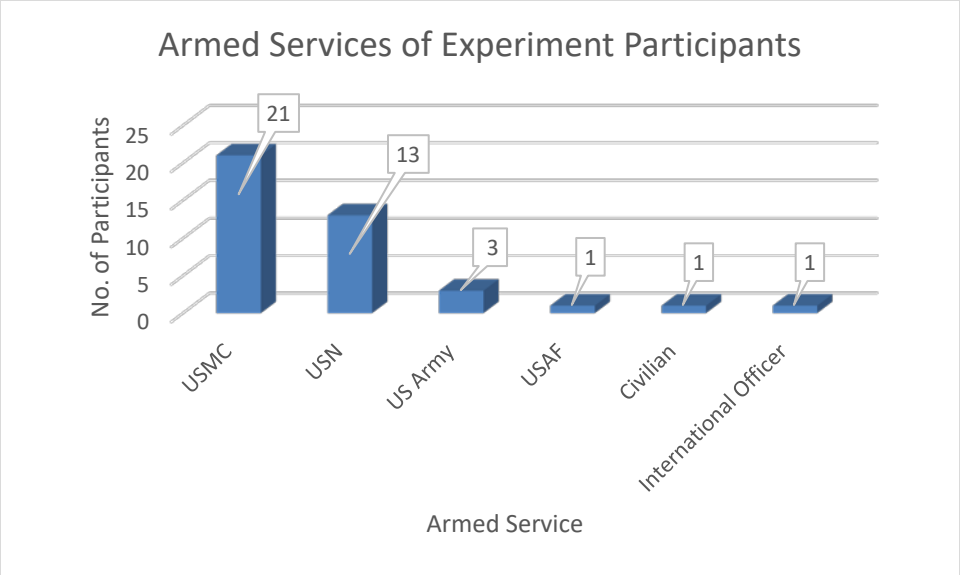


Figure 11. Armed Services of Experiment Participants.

Experiment participants were predominately male with 35 of 40 participants totaling 88% representation. Additionally, the mean age of the volunteers was 33.25 years of age. See Figure 12 for participant age and gender breakdown.

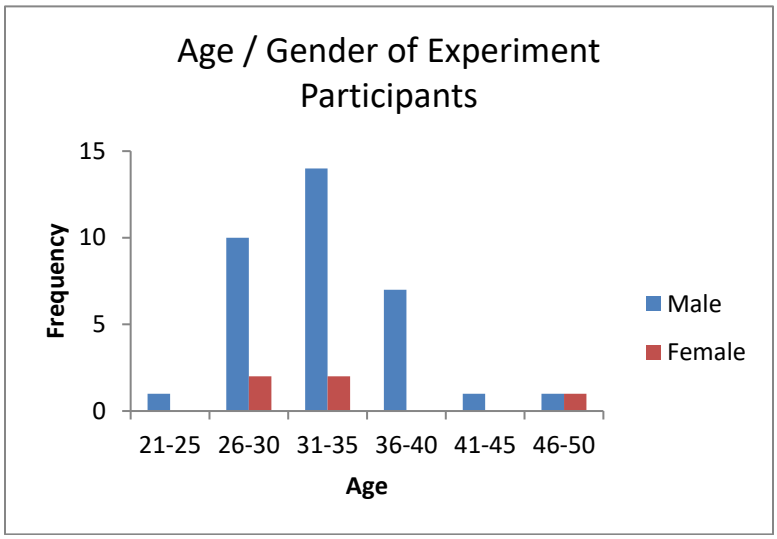


Figure 12. Age and Gender of Experiment Participants.

B. MATERIALS

1. TMPLAR

The most essential material to the conduct of this study was the TMPLAR application itself. The TMPLAR iteration used in this study was web-based and accessible from any web browser using a non-Domain Name Service (DNS)-qualified, public internet protocol (IP) Version 4 (V4) address. The application itself was hosted on a web server administered by and located at the UCONN campus in Storrs, Connecticut. The web-based nature of the application allowed for simultaneous trials to be conducted by several participants at separate workstations within the testing site without degraded capability or performance.

The TMPLAR application used in this study utilized two separate portals, one dedicated for user trials and a separate administrator portal. The administrator portal was used primarily for assigning scenarios to each user profile with approximately equal distributions of low, moderate, and high uncertainty scenarios. Additionally, the researchers had the ability to "Black List," or otherwise remove any scenarios that were technically problematic or faulty from any of the user profiles and replace them with others of equal uncertainty levels. The researchers also were able to use the administrator portal to view the progress of users as they completed their assigned scenarios in real time. Lastly, the administrator portal allowed the researchers to download log files detailing the users' route selections in addition to the timestamps of those selections, confidence ratings, and a record containing X and Y screen coordinates of each mouse click.

The user portal was the mechanism through which both the experiment was conducted and the time to decision (TTD) and route selection data was collected. Like the administrator portal, the user portal was accessed through a public IP address and viewable through any standard commercial internet connection. Volunteer interaction with the user portal was initiated by a splash page, prompting the user to enter a researcher-provided, pseudo-random identification number, as shown in Figure 13.

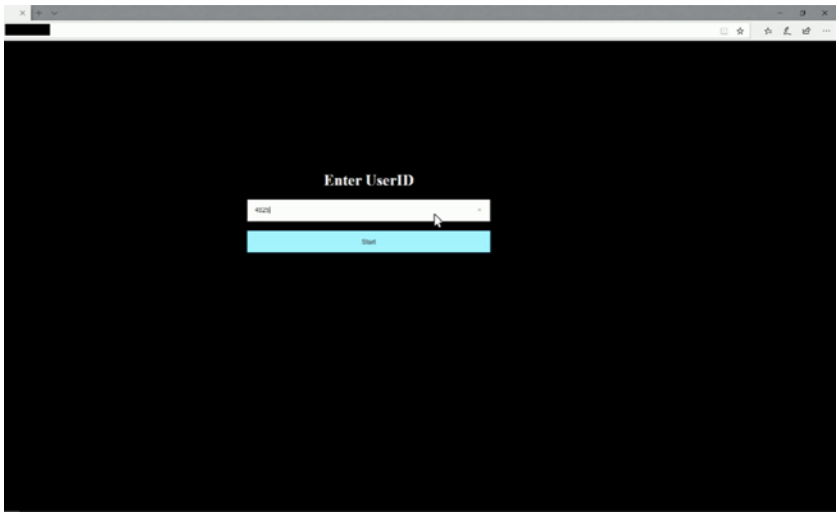
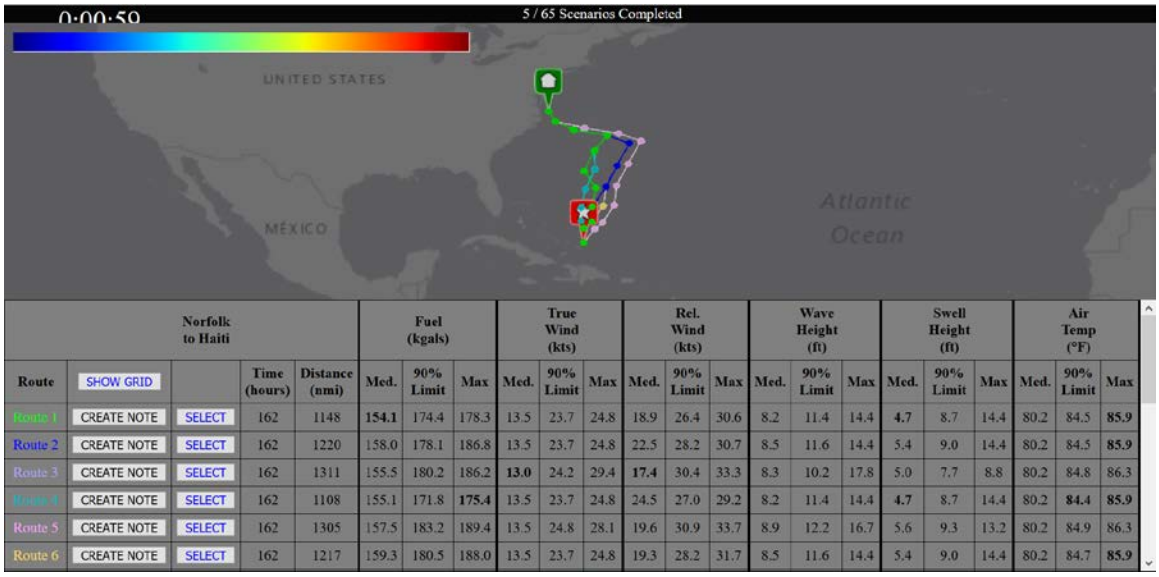


Figure 13. TMPLAR User Splash Page.

The user identification number corresponded to a set of 65 route planning scenarios, with the initial five serving the purpose of orienting the user to the interface and the remaining 60 being recorded for analysis. After the identification number was entered, the user was then taken to one of three DSS interfaces, depending on which DSS level was assigned. Figure 14 illustrates the user interface for the Level 0 DSS.



			Norfolk to Haiti		Fuel (kgals)			True Wind (kts)			Rel. Wind (kts)			Wave Height (ft)			Swell Height (ft)			Air Temp (°F)		
Route	SHOW GRID		Time (hours)	Distance (nmi)	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max
Route 1	CREATE NOTE	SELECT	162	1148	**154.1**	174.4	178.3	13.5	23.7	24.8	18.9	26.4	30.6	8.2	11.4	14.4	**4.7**	8.7	14.4	80.2	84.5	**85.9**
Route 2	CREATE NOTE	SELECT	162	1220	158.0	178.1	186.8	13.5	23.7	24.8	22.5	28.2	30.7	8.5	11.6	14.4	5.4	9.0	14.4	80.2	84.5	**85.9**
Route 3	CREATE NOTE	SELECT	162	1311	155.5	180.2	186.2	**13.0**	24.2	29.4	**17.4**	30.4	33.3	8.3	10.2	17.8	5.0	7.7	8.8	80.2	84.8	86.3
Route 4	CREATE NOTE	SELECT	162	1108	155.1	171.8	**175.4**	13.5	23.7	24.8	24.5	27.0	29.2	8.2	11.4	14.4	**4.7**	8.7	14.4	80.2	**84.4**	**85.9**
Route 5	CREATE NOTE	SELECT	162	1305	157.5	183.2	189.4	13.5	24.8	28.1	19.6	30.9	33.7	8.9	12.2	16.7	5.6	9.3	13.2	80.2	84.9	86.3
Route 6	CREATE NOTE	SELECT	162	1217	159.3	180.5	188.0	13.5	23.7	24.8	19.3	28.2	31.7	8.5	11.6	14.4	5.4	9.0	14.4	80.2	84.7	**85.9**

Figure 14. TMPLAR Level 0 DSS Interface.

The Level 0 DSS is divided into two fields. At the top of the screen is Field I, which is a low-detail map segment that allowed users to pan and zoom along the generated routes with the computer mouse. Field II at the bottom of the screen contained data for relevant route attributes, with bolded figures indicating the most advantageous value for a given attribute. Clicking on values in the, "Route," column populated corresponding color-coded routes in Field I, allowing users to visually compare route tracks. Final route selections for each scenario were made by clicking on the, "Select," Button that corresponded with the desired route. Upon making their selection, users were then prompted to rate their confidence in the chosen route on a 1 – 7 Likert Scale.

Increasing in complexity and the amount of data made available to the user, the Level 1 DSS added AI-augmentation underneath the user interface. A screenshot of the Level 1 DSS interface is shown in Figure 15.

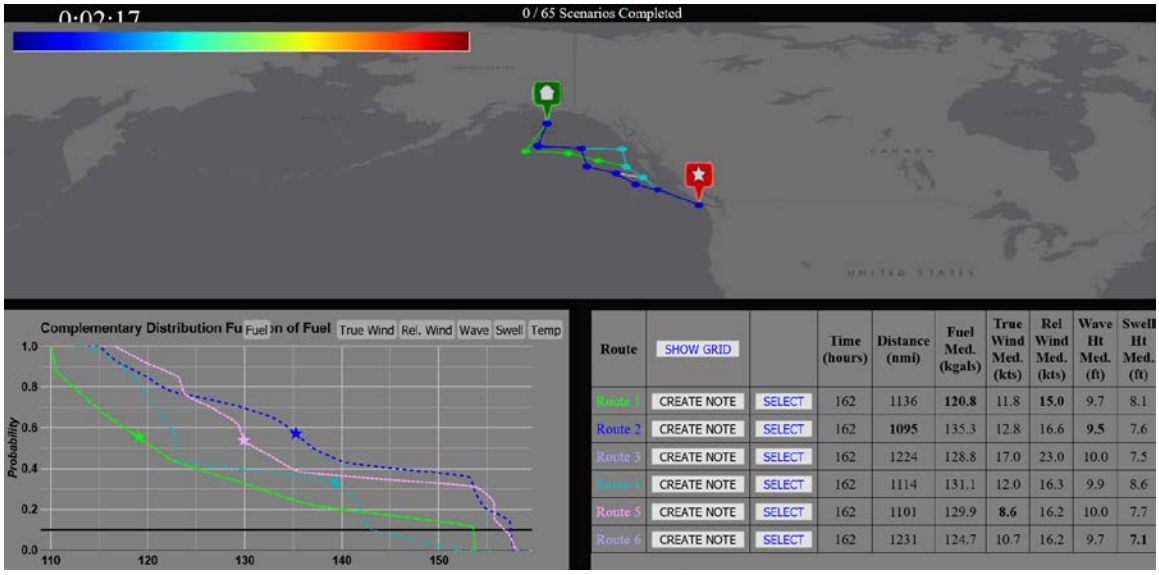


Figure 15. TMPLAR Level 1 DSS Interface.

For the purposes of analysis and standardization, DSS Levels 1 and 3 generally follow field designations as those used in quadrants of a Cartesian Plane. Field I, as in the Level 0 DSS, is still represented by the map at the top of the screen. Field II containing the cumulative distribution function was located at the bottom left of the screen. By selecting routes in the

leftmost column of the data table, probability regressions would be displayed in Field II and route lines populated in Field I, with median values based off of the simulations represented as stars along the Field II regressions. The tabs at the top of Field II allowed uses to switch between route attributes for comparative analysis. Field III was located at the bottom right of the interface and contained the route data table as in Level 0. The procedure for route selection and confidence rating remained the same as in DSS Level 0.

The TMPLAR DSS Level 3 interface, as shown in Figure 16, again incorporated AI-augmentation to its simulation algorithm and provided the user with the most tools pertaining to the required task.

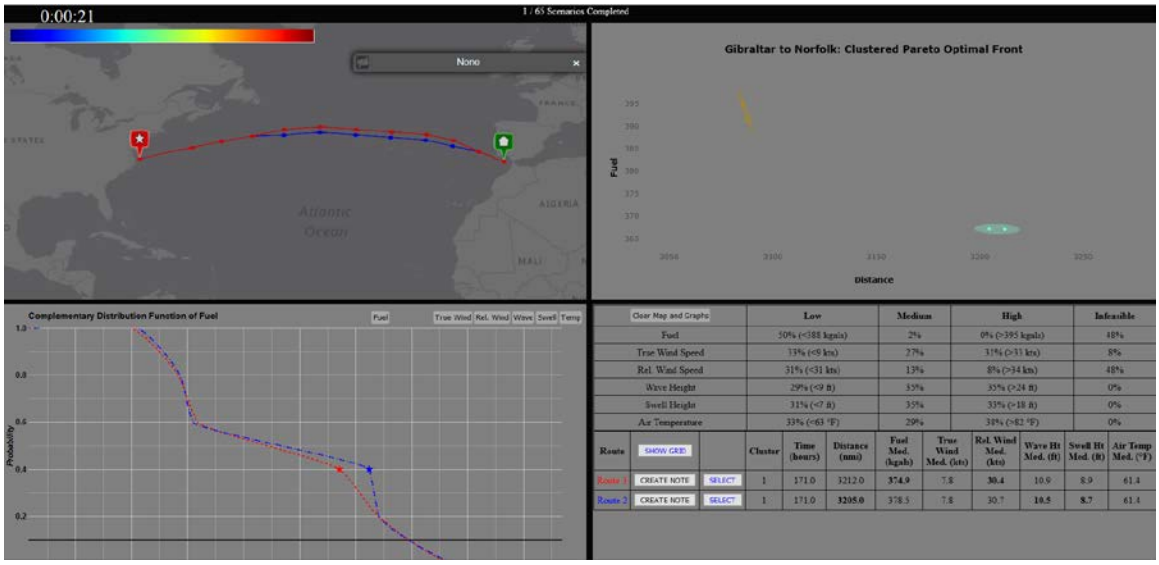


Figure 16. TMPLAR Level 3 DSS Interface.

Beginning with Field I in the top right, the user was presented with a graphical depiction of the Pareto-optimal front weighing fuel usage and distance. By clicking on the ovals grouping together like routes, route attribute data would appear below in Field IV. The next field counterclockwise at the top left was Field II, with the map having the same functionality as in the previous DSS Levels. Field III in the bottom left contained the cumulative distribution function, again with functionality as before. Field IV at the bottom right contained the route attribute data table. Again, clicking on the values in the leftmost

column populated the map with graphical route tracks and the cumulative distribution function with probability regressions.

2. Eye Tracking / Pupillometry

Eye tracking and pupillometry data was captured with the Gazepoint suite, which included the GP3HD tracker, Gazepoint Analysis UX Edition v6.0.0, and Gazepoint Control x64 applications. The Gazepoint suite allows users to calibrate the eye tracker to an experiment participant, build a customized experiment, and record/analyze an experiment's results (Vogl, 2014). User-defined/customizable eye tracking projects that can be conducted with the Gazepoint suite include general screen capture, projects involving text, images, or video, and web interface navigation (Gazepoint, 2019). The GP3HD eye tracker model used in this experiment had a user-selectable sample rate of 60/150Hz and accuracy of gaze measurement of 0.5 – 1^{o} from the user focal point (Gazepoint 2020). Gazepoint Analysis supported a five or nine-point user calibration and had a field of view of 35cm x 22 cm with a plus or minus 15cm range of depth movement (Gazepoint 2020). The GP3HD tracker itself measured 320 x 45 x 47 mm in dimensions and had a mass of 155g (Gazepoint 2020). The tracker was mounted directly beneath the computer screen where participants conducted the experiment as shown in Figure 17.

Figure 17. Participant Workstation.

A Gazepoint Analysis user-defined profile was configured specifically for this experiment. This was done first by launching the Gazepoint Analysis application. Upon loading, Gazepoint Analysis prompted the researcher to open a saved project or begin a new project. After electing to begin a new project, Gazepoint Analysis then prompted the researcher to select the media to be used for this experiment. As this experiment used a web-based application, the IP address of the TMPLAR application was entered and saved. After this was accomplished, the researcher was given access to the Gazepoint Analysis user interface, as shown in Figure 18.

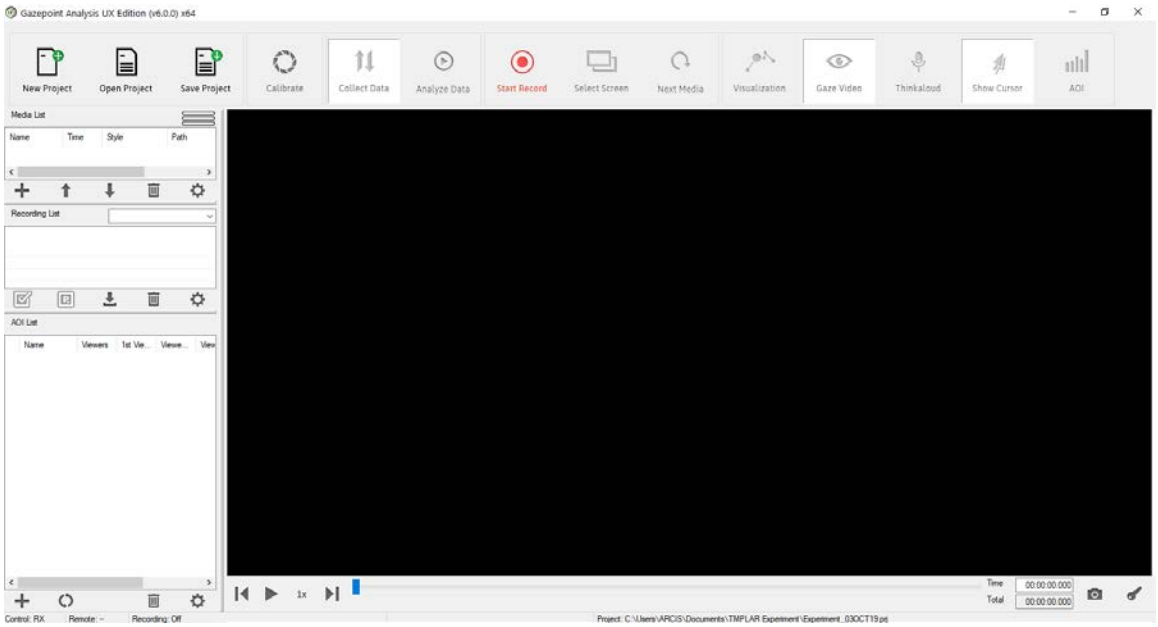


Figure 18. Gazepoint Analysis User Interface.

The left side of the user interface allows the researcher to select between media used for a recording, such as a different video or website for different participants, select or view previously made recordings, or establish Areas of Interest (AOI) for analysis and assign then to user profiles. The ribbon near the top of the screen contained basic functions, allowing the researcher to load and save a project, calibrate the GP3HD sensor to a user, begin a recording, select the active screen for multi-screen displays, and display researcher-defined AOI zones. After the project shell was created and the TMPLAR application IP address loaded in the media list, Gazepoint Analysis was configured to record and collect basic eye tracking and pupillometry data for users as they conduct the experiment.

The Gazepoint Control application was the final major component of the Gazepoint suite that was utilized for this study. Gazepoint Control acts as a, “Power Switch,” for the GP3HD tracker and needs to be running in order to collect data (Vogl, 2014). The functions performed by the Gazepoint Control application are largely automatic and require no action from the researcher (Vogl, 2014). For the purposes of this study, Gazepoint Control was almost exclusively used to calibrate the eye tracker to each participant before the

experiment trials began. See Figure 19 for a graphic depicting the Gazepoint Control interface.

Figure 19. Gazepoint Control User Interface.

The ribbon at the top of the window in the Gazepoint Control interface contained the relevant commands for the researcher. Of most importance, "Calibrate," launched the calibration sequence for each user. Other options including sampling rate were left as default, and biometrics, such as heart rate monitoring, were supported by Gazepoint but not utilized for this experiment.

3. Exit Questionnaire

The final material used to collect data was an exit questionnaire administered to all participants. The questionnaire used a Likert Scale for user ordinal ranking of efficiency of the TMPLAR application, its overall utility, and to rate their overall experience in maritime route planning. Fill in the blank responses were recorded capturing demographic data including age, gender, hours of sleep the previous night, etc. Finally, narrative responses were recorded for user suggestions for improvement to the TMPLAR application and other

maritime route planning tools used, if applicable. The Exit Questionnaire instrument is shown in full in Appendix A.

C. PROCEDURE

1. Pilot Test

A pilot test of the TMPLAR application was conducted from 26 AUG – 13 SEP 2019 aboard the NPS Campus in Glasgow Hall. The goal of the pilot test was to trail the TMPLAR application and observe proper function, determine if any scenarios did not function correctly and needed to be added to the, "Black List," refine the participant in-brief and exit questionnaire, and determine the approximate time required for each participant to conduct the experiment. Pilot test volunteers were recruited by word of mouth, with 12 personnel participating. Upon arrival to the testing space, participants were immediately given an Informed Consent document, which they duly signed. See Appendix B for the Informed Consent form. After the Informed Consent form was signed by the participant, the student researcher explained the function of the TMPLAR application, the goals of the pilot test and the study, and the specific tasks assigned to the participant. After the verbal brief was complete and all questions were answered, the participants were assigned a user identification number, which was then entered in the TMPLAR splash page to begin the assessment.

Of the 12 participants for the pilot test, three personnel each were assigned to the DSS 0, DSS 1, and DSS 3 interfaces. Each participant completed 10 scenarios across all uncertainty levels in their assigned DSS. The progress of the participants was closely monitored for any issues with the TMPLAR application. After the participants completed their scenarios, the Exit Questionnaire was administered. When the Exit Questionnaire was complete, the student researcher engaged the participants in a verbal debrief to discuss functionality of TMPLAR, clarity of instructions and assigned tasks, clarity of Exit Questionnaire, and any proposed points for improvement. From the pilot test, it was determined that full 65 scenario trials would likely take between 60–90 minutes to complete and that full user manual-style written instructions were necessary.

2. Experimentation

The main experiment was conducted between the period of 28 OCT 2019 – 28 FEB 2020 in Glasgow Hall, Room 103. A total of 40 volunteers participated in the main experiment. Upon sign up, the student researcher contacted the volunteer via email to confirm the time and send electronic copies of the Informed Consent Form and the TMPLAR User Manual for their assigned DSS level. Before participant arrival to the experiment, the student researcher checked the GP3HD tracker, the Gazepoint Analysis and Gazepoint Control Applications, and TMPLAR for correct function to ensure that the experiment trails occurred without issue.

Upon arrival of the volunteers to the research space, they were immediately greeted and presented with the Informed Consent form, which was then signed and filed. Next, the participants received a verbal brief from the student researcher covering the goals of the study and the tasks assigned to the participants. The verbal brief is enclosed in Appendix C. They were then presented with a hard copy of the TMPLAR User Manual (DSS 0: Appendix D, DSS 1: Appendix E, DSS 3: Appendix F), which the student researcher then used to talk through the layout of the TMPLAR interface and the functions available to the participant. Participants were also given, “TMPLAR Quick Start,” graphics depicting the steps required to complete a scenario in their given TMPLAR interface as shown in Appendix G. Once the instructions and experiment goals were understood, the student researcher and participant moved on to eye tracker calibration.

Eye tracker calibration was performed through the Gazepoint Control application and conducted for every user. The participant was instructed to center him or herself in front of the computer monitor and assume a posture that was natural to them and comfortable. The student researcher then examined the Gazepoint Control interface to ensure that the user was centered in the sensor’s field of view and that the GP3HD device was tracking their eyes. If the user was not centered or the GP3HD device could not track the user’s eyes, body position and sensor orientation were adjusted until the desired result was achieved. The calibration sequence was then initiated, in which a white dot appeared on a black background on the display and moved to the four corners, then ending in center of the monitor. Gazepoint Control then provided a prompt notifying the student researcher

that calibration was either successful or that calibration had failed for one or both eyes. If calibration had failed for a particular user, repeated attempts for successful calibration were made for as long as practical without jeopardizing the timeline needed for the volunteer to conduct the experiment.

After the calibration sequence, users completed five practice scenarios, in which they were encouraged to explore the user interface and ask the student researcher questions when required. This ensured that the users were comfortable with TMPLAR and that they understood all of the application's functions that were made available to them. After the five practice scenarios were completed, the participants seamlessly transitioned to 60 route planning scenarios that were recorded as participant results. While the participants worked their way through the scenarios, the student researcher observed their progress through the Gazepoint Analysis window on his own separate display, as shown in Figure 20.

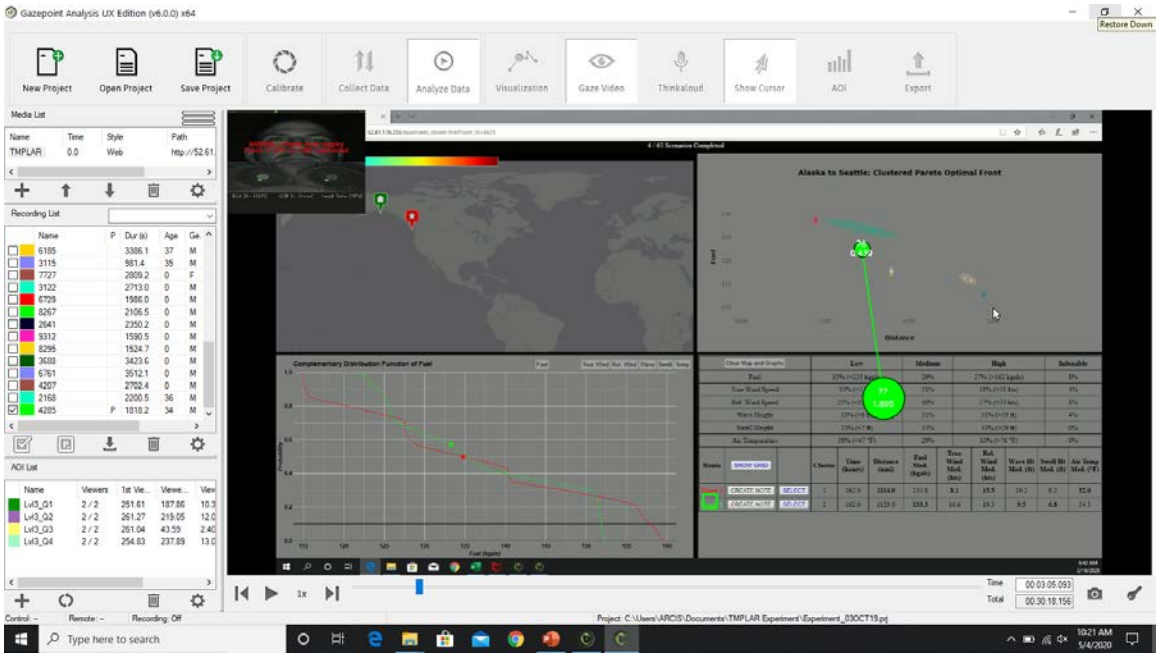


Figure 20. Experiment In-Progress View through Gazepoint Analysis Interface

Through the in-progress view in the Gazepoint Analysis window, the student researcher was not only able to determine the user's relative progress at any point in the

experiment, but was able to give them corrective verbal cues if their posture was affecting the GP3HD sensor's ability to measure their eye movements. Upon completion of the last route planning scenario, TMPLAR prompted the users notifying them that the experiment was complete. The recording was stopped and saved in Gazepoint Analysis and the participants were given the exit questionnaire. After completing the exit questionnaire, participants were then verbally debriefed and their responses captured.

3. Data Retrieval and Structuring

After the departure of the volunteer, the student researcher navigated to the TMPLAR administrator portal and downloaded the aggregated dump file for the experiment trials. The dump.zip file contained five separate .csv files, which captured participant performance during the experiment. The first of these was named, "notes.csv," and captured any entries made by participants using the notes function in TMPLAR. The, "timer.csv," file captured total elapsed times for experiment completion across users. The third file in the dump,zip pack was, "scenarios.zip," and included the numeric codes for all scenarios assigned to each User Identification number, totaling 65 per user. The, "chosenroutes.csv," file was of high utility in that it contained 65 entries per user, recording the User Identification, scenario number, route selected, user confidence rating of their selection, and the timestamp of their selection. The final file made available through the TMPLAR administrator portal was entitled, "logdata.csv." This file contained User Identification numbers, X and Y screen coordinates with descriptions of screen elements that users clicked, and corresponding timestamps. Data from these files was screened for validity, duplicate values were identified and omitted when required, and data was transferred to individual and collective spreadsheets corresponding to assigned DSS level. Additionally, Exit Questionnaire responses were transferred to spreadsheets for later analysis.

Raw click data, as previously discussed, was provided in the, "logdata.csv," dump file. Since this file provided X and Y screen coordinates logging every mouse click made by the user, a pixel map of the display was used to categorize clicks into fields corresponding with the DSS-unique divisions of the interface. Using a screen resolution of

1920 x 1080 in the experiment's display, the pixel maps displayed in Figure 21 were used to categorize user clicks recorded by TMPLAR.

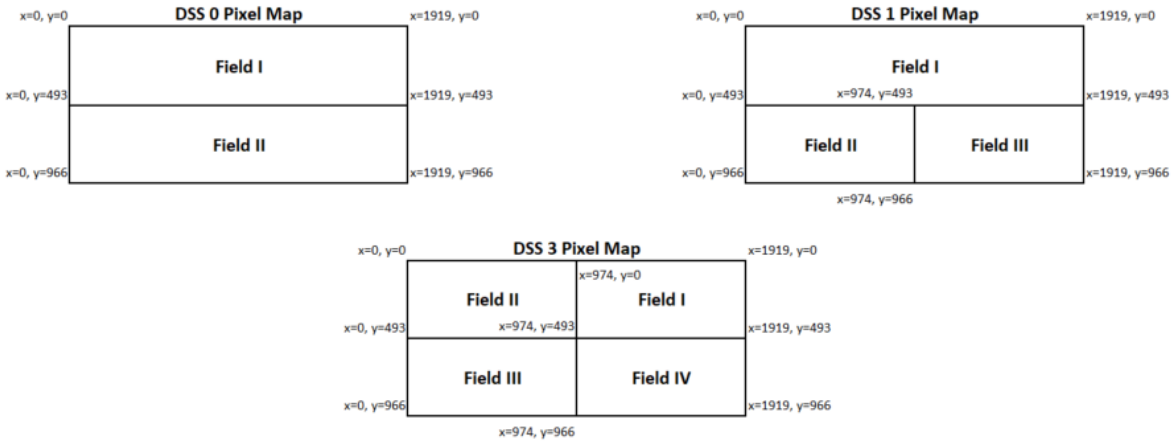


Figure 21. Pixel Maps Used for User Click Categorization.

In addition to recording frequency of user clicks in specific fields, the, "logdata.csv," file also noted when users made their route selection, bringing each scenario to an end and progressing to the next. The notation of, "Select Route," in the log file allowed for fault checking for both time to complete each scenario and the completion of all scenarios for each user.

The retrieval of user eye fixation and tracking data was initiated through the Gazepoint Analysis Interface. First, a user's profile was selected under the, "Recording List," header and the video recording slider at the bottom of the screen was moved to the beginning of the sixth scenario. This was done to discount the first five scenarios that were dedicated to practice only. Next, the "+," button was selected beneath the, "AOI List," pane. Once this was done, the mouse cursor changed from a standard cursor to a crosshair pattern, allowing the student researcher to designate an AOI by freehand. The freehand option was not used because the greatest amount of consistency between users was desired. Instead, a general square pattern was made in the approximate area that the AOI was to be placed, then the AOI was double-clicked in the, "AOI List," pane. A prompt then appeared,

allowing for the manual entry of X and Y coordinates for the AOI's origin and width and height dimensions for the AOI itself. Refer to Figure 22 for AOI dimensions used for each level of DSS.

DSS 0 AOI Map

Field I	
X	0
Y	0.068
Width	1
Height	0.45
Field II	
X	0
Y	0.528
Width	1
Height	0.435

DSS 1 AOI Map

Field I		
X	0	
Y	0.068	
Width	1	
Height	0.461	
Field II		**Field III**
X	0	X 0.5
Y	0.528	Y 0.528
Width	0.5	Width 0.5
Height	0.435	Height 0.435

DSS 3 AOI Map

Field II		Field I	
X	0	X	0.508
Y	0.068	Y	0.068
Width	0.508	Width	0.492
Height	0.458	Height	0.458
Field III		**Field IV**	
X	0	X	0.508
Y	0.523	Y	0.523
Width	0.508	Width	0.492
Height	0.446	Height	0.446

Figure 22. Pixel Maps Used for AOI Designation.

This process was repeated for each AOI required the user's assigned DSS level. When all AOIs for a given user were designated, the video recording slider was advanced to the completion of the 65th scenario and paused. The, "Calculate," button under the, "AOI List," captured the percentages of users gazes in each AOI. This data was then exported, and the resulting file contained the time each AOI was viewed in seconds, time each AOI was viewed as a percentage, number of visual fixations per AOI, and number of visual revisits that a user made to an AOI. Unlike the process used to categorize mouse clicks into fields where the same pixel map was applied for all users, each AOI was entered manually for each user to account for the varying times that each user began scenario six and finished scenario 65.

The process used to retrieve the pupillometry data collected through the Gazepoint suite was comparatively straightforward. The user profile was selected and the, "Export," button was clicked as before. Here, all data options were selected for export. See Appendix H for a complete listing of Gazepoint Analysis output data fields. Once the .csv file was

exported, it was organized into user-specific folders along with the TMPLAR dump files, click analysis, and AOI/eye tracking data for further analysis.

IV. RESULTS

A. TIME TO DECISION

User time to decision was recorded in TMPLAR and made available in a .csv file for analysis. These data were organized into the three levels of DSS users and three different scenario difficulties. Additionally, one participant was omitted from the analyses due to the user not completing all scenarios assigned. User time to decision data was first grouped into the DSS level assigned to the user. After this was done, a 3 (DSS level) x 3 (scenario difficulty) repeated measures ANOVA was conducted to determine what, if any significant differences were present in user time to decision. After analysis, it was apparent that there was a significant main effect of decision time across the DSS levels; $F(2, 36) = 3.31$, $p = 0.05$, and $\eta_p^2 = 0.16$. Additionally, a significant main effect of decision time across the difficulty, $F(2, 117) = 5.63$, $p = 0.01$, and $\eta_p^2 = 0.09$ was present. There was no interaction between DSS level and difficulty, $p = 0.52$.

Follow-up t-tests for the main effect of DSS revealed a significant difference between DSS 0 and DSS 3, $t(24) = -2.03$, $p = 0.05$, and DSS 1 and DSS 3, $t(23) = -2.80$, $p = 0.01$. There was no difference between DSS 0 and DSS 1, $p = 0.28$. Participants in the DSS 3 condition took longer to make a decision than those in DSS 0 and DSS 1 conditions. However, there was no difference in decision time between the DSS 0 and DSS 1 conditions. Descriptive statistics and pairwise comparisons across DSS levels are illustrated in Table 1.

Table 1. User Time to Decision across DSS Levels

User TTD Across DSS

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	14	30.18	11.47	3.07	23.56	36.80
DSS 1	13	25.82	13.51	3.75	17.66	33.98
DSS 3	12	39.44	15.61	4.51	29.52	49.36
Pairwise Comparisons						
DSS A	DSS B	Mean Difference (A - B)	df	Std. Error	Sig.	t
DSS 0	DSS 1	0.10	25	0.08	0.28	1.10
DSS 0	DSS 3	-0.11	24	0.08	0.05	-2.03
DSS 1	DSS 3	-0.21	23	0.08	0.01	-2.80

T-tests for scenario difficulty revealed significant differences between Easy and Hard, $t(78) = -3.35$, $p < 0.01$ scenarios. There was no difference between Easy and Medium, $p = 0.09$ and Medium and Hard, $p = 0.11$ scenarios. Participants making route selections classified as Hard took longer to make a decision than those classified as Easy. Descriptive statistics and pairwise comparisons across difficulty conditions are in Table 2.

Table 2. User Time to Decision across Difficulty

User TTD Across Difficulty

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Easy	14	24.70	10.57	1.69	21.27	28.13
Medium	13	30.26	13.23	3.75	25.97	34.55
Hard	12	38.79	19.85	2.12	32.35	45.22
Pairwise Comparisons						
Difficulty A	Difficulty B	Mean Difference (A - B)	df	Std. Error	Sig.	t
Easy	Medium	-0.09	78	0.01	0.09	-1.73
Easy	Hard	-0.18	78	0.01	0.00	-3.35
Medium	Hard	-0.09	78	0.01	0.11	-1.63

B. USER CLICKS WITHIN TMPLAR INTERFACE

User interaction with the TMPLAR interface was quantified through the mapping and recording of mouse clicks. Level of user interaction with the TMPLAR interface fields was assessed to determine (a) what data and presentation mode was most utilized by users to make route selections, and (b) if the reliance on data presented in screen fields common across all DSS levels changed as more data was made available to the user. To assess what data field was most utilized by users within DSS levels, measurements reporting clicks in each screen field were divided by overall click counts for each user, yielding percentages by field. Click data was then compiled, extreme outliers were identified and omitted, and the data set was transformed using a $LOG_{10}(x + 1)$ algorithm due to non-normal distributions identified by abnormal skewness and kurtosis.

A univariate ANOVA conducted on DSS 0 user mouse clicks revealed a significant main effect of screen field within DSS 0; $F(1, 26) = 35.44$, $p < 0.01$, and $\eta_p^2 = 0.58$. Therefore, it was determined that users interacted with the Table through mouse clicks at statistically significant higher rates than the Map. Refer to Table 3 for descriptive statistics of DSS 0 click data.

Table 3. Analysis of DSS 0 Click Data.

Percentage of User Clicks by Screen Field: DSS 0

Descriptive Statistics						
Field	N	Mean %	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Map	14	0.16	0.13	0.03	0.09	0.24
Table	14	0.84	0.13	0.03	0.76	0.91

DSS 1 click data were analyzed using a univariate ANOVA with allowances made for the inclusion of the Cumulative Distribution Function (CDF) screen field. The ANOVA revealed a significant main effect of screen field, $F(2, 27) = 253.94$, $p < 0.01$, and $\eta_p^2 = 0.95$. Follow-on t-tests confirmed significant differences between the Map and CDF, $t(18) = 10.59$, $p < 0.01$, Map and Table, , $t(18) = -7.93$, $p < 0.01$, and the CDF and Table fields, $t(18) = -102.04$, $p < 0.01$. From this analysis, it was determined that DSS 1 users clicked

the Table field at significantly higher rates than either the Map or the CDF. Also, users interacted with the Map screen field at significantly higher rates than the CDF field. Refer to Table 4 for descriptive statistics and pairwise comparison of DSS 1 click data.

Table 4. Analysis of DSS 1 Click Data.

Percentage of User Clicks by Screen Field: DSS 1

Descriptive Statistics						
Field	**N**	**Mean %**	**Standard Deviation**	**Standard Error**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
Map	10	0.15	0.11	0.04	0.07	0.23
CDF	10	0.00	0.00	0.00	0.00	0.00
Table	10	0.85	0.11	0.04	0.77	0.93
Pairwise Comparisons (Transformed Data)						
Field A	**Field B**	**Mean Difference (A - B)**	**df**	**Std. Error**	**Sig.**	**t**
Map	CDF	1.10	18	0.09	0.00	10.59
Map	Table	-0.84	18	0.09	0.00	-7.93
CDF	Table	-1.932	18	0.09	0.00	-102.04

A univariate ANOVA was then conducted for DSS 3 click data with the inclusion of the Pareto Distribution screen field. The ANOVA reported a significant main effect of screen field, $F(3, 28) = 144.38$, $p < 0.01$, and $\eta_p^2 = 0.94$. The results of t-tests reported significant differences between the Pareto Distribution and Map, $t(14) = 8.74$, $p < 0.01$, Pareto Distribution and CDF, $t(14) = 45.22$, $p < 0.01$, Pareto Distribution and the Table, $t(14) = -5.19$, $p < 0.01$, the Map and Table, $t(14) = -10.32$, $p < 0.01$, and the CDF and Table fields, $t(14) = -96.09$, $p < 0.01$. The results show that DSS 3 users, through mouse clicks, interacted with the Pareto Distribution and the Table significantly more than the CDF or the Map. Additionally, a significant difference was determined to be present between the Pareto Distribution and the Table. Refer to Table 5 for descriptive statistics and pairwise comparison of DSS 3 click data.

Table 5. Analysis of DSS 3 Click Data.

Percentage of User Clicks by Screen Field: DSS 3

Descriptive Statistics						
Field	N	Mean %	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Pareto	8	0.38	0.09	0.03	0.30	0.45
Map	8	0.02	0.03	0.01	0.00	0.05
CDF	8	0.00	0.00	0.00	0.00	0.00
Table	8	0.60	0.07	0.03	0.54	0.66
Pairwise Comparisons (Transformed Data)						
Field A	Field B	Mean Difference (A - B)	df	Std. Error	Sig.	t
Pareto	Map	1.30	14	0.11	0.00	8.74
Pareto	CDF	1.58	14	0.11	0.00	45.22
Pareto	Table	-0.21	14	0.11	0.00	-5.19
Map	CDF	0.28	14	0.11	0.08	1.92
Map	Table	-1.51	14	0.11	0.00	-10.32
CDF	Table	-1.784	14	0.11	0.00	-96.09

To determine the screen fields most utilized by users across DSS levels, screen elements common across all DSS levels, which included the Map and Table fields, were examined. The univariate ANOVA conducted using pooled Map and Table data for all DSS levels revealed a significant main effect of screen field; $F(1, 26) = 104.69$, $p < 0.01$, and $\eta_p^2 = 0.63$. As a result, it was determined that users interacted with the Table through mouse clicks at statistically significant higher rates than the Map across all DSS levels. Refer to Table 6 for descriptive statistics of the common screen elements click data.

Table 6. Analysis of Common Screen Elements Click Data.

Percentage of User Clicks by Screen Field: Common Screen Elements

Descriptive Statistics						
Field	N	Mean %	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Map	32	0.12	0.12	0.02	0.08	0.17
Table	32	0.78	0.15	0.03	0.73	0.84

C. VISUAL FIXATIONS

As previously discussed, user gaze fixations were recorded using the Gazepoint software suite and the GP3HD eye tracker. As a result of inconsistent sampling for each user caused by the different amounts of time used by each participant, percentages of total recorded gazes per each screen field were used to feed the following analysis. As before, extreme outliers were identified and omitted. The data set was transformed using a $LOG_{10}(x)$ algorithm due to non-normal distributions. As for the click data, this transformed data set was then used to determine if any statistically significant difference existed between user gaze rates for screen elements (map, data table, cumulative distribution function, Pareto distribution) within DSS levels and for screen elements common across DSS levels (map and table only). The comparative rates of gaze fixations were assessed to determine (a) what data and presentation mode did users fixate upon the most as they made route selections, and (b) if the reliance on data presented in screen fields common across all DSS levels changed as more data was made available to the user.

A Univariate ANOVA performed DSS 0 visual fixations revealed a significant main effect of screen field within DSS 0; F (1, 20) = 58.51, $p < 0.01$, and $\eta_p^2 = 0.575$. From this information, it was determined that users visually fixated on the Table field at statistically significant higher rates than the Map. Refer to Table 7 for descriptive statistics of DSS 0 visual fixations data.

Table 7. Analysis of DSS 0 Visual Fixations Data.

Percentage of Fixations by Screen Field: DSS 0

Descriptive Statistics						
Field	N	Mean %	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Map	11	0.22	0.14	0.04	0.12	0.32
Table	11	0.78	0.14	0.04	0.68	0.88

The analysis continued with DSS 1 fixations data, with the ANOVA reporting a significant main effect of screen field, F(2, 30) = 18.16, $p < 0.01$, and $\eta_p^2 = 0.55$. T-tests revealed significant differences between the Map and Table, t(20) = -6.15, $p < 0.01$, and

the CDF and Table fields, $t(20) = -5.68$, $p < 0.01$. From these figures, it was determined that DSS 1 users visually fixated on the Table field at significantly higher rates than either the Map or the CDF, for which there was no significant difference between. Refer to Table 8 for descriptive statistics and pairwise comparison of DSS 1 visual fixations data.

Table 8. Analysis of DSS 1 Visual Fixation Data.

Percentage of Fixations by Screen Field: DSS 1

Descriptive Statistics						
Field	**N**	**Mean %**	**Standard Deviation**	**Standard Error**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
Map	11	0.18	0.17	0.05	0.06	0.29
CDF	11	0.14	0.12	0.04	0.06	0.22
Table	11	0.68	0.18	0.05	0.57	0.80
Pairwise Comparisons (Transformed Data)						
Field A	**Field B**	**Mean Difference (A - B)**	**df**	**Std. Error**	**Sig.**	**t**
Map	CDF	0.17	20	0.16	0.39	0.88
Map	Table	-0.72	20	0.16	0.00	-6.15
CDF	Table	-0.890	20	0.16	0.00	-5.68

The Univariate ANOVA was repeated for DSS 3. Once again, the ANOVA reported a significant main effect of screen field, $F(3, 36) = 30.49$, $p < 0.01$, and $\eta_p^2 = 0.72$. Individual t-tests described significant differences between the Pareto Distribution and Map, $t(18) = 7.23$, $p < 0.01$, Pareto Distribution and CDF, $t(18) = 3.84$, $p < 0.01$, Pareto Distribution and Table, $t(18) = -3.10$, $p = 0.01$, the Map and Table, $t(18) = -9.63$, $p < 0.01$, and the CDF and Table fields, $t(18) = -5.65$, $p < 0.01$. From this information, it was determined that DSS 3 users visually fixated on the Pareto Distribution and the Table with significantly higher frequency than the CDF or the Map. As in the results from the DSS 3 click data, no significant differences were shown to exist between the CDF and Map fields. Refer to Table 9 for descriptive statistics and pairwise comparison of DSS 3 visual fixation data.

Table 9. Analysis of DSS 3 Visual Fixations Data.

Percentage of Fixations by Screen Field: DSS 3

Descriptive Statistics						
Field	N	Mean %	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Pareto	10	0.32	0.12	0.04	0.24	0.41
Map	10	0.06	0.04	0.01	0.03	0.09
CDF	10	0.12	0.07	0.02	0.08	0.17
Table	10	0.50	0.12	0.04	0.41	0.58

Pairwise Comparisons (Transformed Data)						
Field A	Field B	Mean Difference (A - B)	df	Std. Error	Sig.	t
Pareto	Map	0.80	18	0.12	0.00	7.23
Pareto	CDF	0.49	18	0.12	0.00	3.84
Pareto	Table	-0.20	18	0.12	0.01	-3.10
Map	CDF	-0.31	18	0.12	0.06	-2.01
Map	Table	-1.00	18	0.12	0.00	-9.63
CDF	Table	-0.695	18	0.12	0.00	-5.65

Like in the case of user clicks, screen elements common across all DSS levels were studied to determine which common field (Map or Table) did users visually fixate on the greatest. The univariate ANOVA ran using the aggregated Map and Table data for all DSS levels revealed a significant main effect; $F(1, 62) = 110.17$, $p < 0.01$, and $\eta_p^2 = 0.64$. Mirroring the results from the common screen element user click analysis, users were shown to visually fixate on the Table field at statistically significant higher rates than the Map across all DSS levels. Refer to Table 10 for descriptive statistics of common screen elements visual fixation data.

Table 10. Analysis of Common Screen Elements Visual Fixations Data.

Percentage of Fixations by Screen Field: Common Screen Elements

Descriptive Statistics						
Field	N	Mean %	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Map	32	0.16	0.15	0.03	0.10	0.21
Table	32	0.66	0.19	0.03	0.59	0.72

D. PUPILLOMETRY

Like with user gaze fixations, pupillometry measurements were recorded using the Gazepoint software suite and the GP3HD eye tracker. However, screen fields were not taken into account when pupil diameter was recorded. As a result, raw pupil diameter observations were utilized to describe the effects of DSS level on the relative cognitive load of users. No transformations were applied to the data because of its inherent normal distribution within DSS level. A univariate ANOVA conducted using pupil measurements revealed a significant main effect of screen field across DSS level; $F(2, 103358) = 6192.69$, $p < 0.01$, and $\eta_p^2 = 0.11$. Individual t-tests reported significant differences between the DSS 0 and DSS 1, $t(60014) = 58.71$, $p < 0.01$, DSS 0 and DSS 3, $t(69530) = -42.35$, $p < 0.01$, and DSS 1 and DSS 3, $t(77172) = -111.86$, $p < 0.01$. As a result, it was determined that DSS 3 users expressed significantly greater pupil diameter than DSS 0 users, who in turn showed significantly greater pupil diameter than DSS 1 users. Refer to Table 11 for descriptive statistics and pairwise comparison for user pupillometry data.

Table 11. Analysis of User Pupil Diameter across DSS Levels.

User Pupil Diameter Across DSS

Descriptive Statistics						
Field	N	Mean (mm)	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	26187	3.45	0.49	0.00	3.44	3.45
DSS 1	33829	3.22	0.44	0.00	3.22	3.23
DSS 3	43345	3.61	0.51	0.00	3.61	3.62
Pairwise Comparisons						
DSS A	DSS B	Mean Difference (A - B)	df	Std. Error	Sig.	t
DSS 0	DSS 1	0.22	60014	0.00	0.00	58.71
DSS 0	DSS 3	-0.17	69530	0.00	0.00	-42.35
DSS 1	DSS 3	-0.39	77172	0.00	0.00	-111.86

E. CONFIDENCE RATINGS

User confidence in each scenario route selection was recorded in TMPLAR and made available in a .csv file for analysis. To minimize the effects of abnormal skewness and kurtosis, a special data transformation was made prior to statistical analysis. First, user

confidence ratings made on the 1 - 7 Likert Scale were transformed using the reciprocal of the user entry. For example, a user entry of, "1," was transformed to, "7," and an entry of, "2," was transformed to, "6," and so on. The reciprocal of the Likert Scale mitigated the effects of overly negative skewness of the data without compromising the validity of the 3 (DSS level) x 3 (scenario difficulty) repeated measures ANOVA. After the reciprocal was taken, the data was smoothed using the $LOG_{10}(x)$ algorithm to achieve a more normal distribution. From the resulting data set, user confidence across DSS levels and scenario difficulty were analyzed to determine if any significant difference was present. The 3 x 3 repeated measures ANOVA showed that there was no significant main effect of confidence across the DSS levels; $F(2, 36) = 0.46$, $p = 0.64$, and $\eta_p^2 = 0.03$. However, a significant main effect of confidence across difficulty; $F(2, 36) = 73.50$, $p < 0.01$, and $\eta_p^2 = 0.67$, was present. There was no interaction between DSS level and difficulty for user confidence ratings, $p = 0.83$. Descriptive statistics pertaining to mean confidence ratings across DSS levels are illustrated in Table 12.

Table 12. Analysis of User Mean Route Confidence Ratings across DSS Level.

User Mean Confidence Ratings Across DSS

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	14	5.97	1.21	0.32	5.27	6.67
DSS 1	13	5.79	0.83	0.23	5.29	6.29
DSS 3	12	5.92	0.61	0.18	5.53	6.30

Further t-tests applied toward scenario difficulty uncovered significant differences between Easy and Hard, $t(76) = -1.22$, $p < 0.01$ and Medium and Hard, $t(76) = -0.77$, $p < 0.01$, scenarios. From this, it was determined that participants making route selections classified as Hard had higher confidence in their decisions for scenarios classified as Medium or Easy. Users did not have any difference in their confidence ratings for Easy and Medium scenarios. Descriptive statistics and pairwise comparisons for user confidence ratings across difficulty conditions are shown in Table 13.

Table 13. Analysis of User Mean Route Confidence Ratings across Difficulty.

User Mean Confidence Ratings Across Difficulty

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Easy	39	6.05	0.77	0.12	5.80	6.30
Medium	39	5.90	0.94	0.15	5.60	6.21
Hard	39	5.74	1.09	0.17	5.38	6.09
Pairwise Comparisons (Transformed Data)						
Difficulty A	Difficulty B	Mean Difference (A - B)	df	Std. Error	Sig.	t
Easy	Medium	-0.02	76	0.01	0.08	-0.44
Easy	Hard	-0.05	76	0.01	0.00	-1.22
Medium	Hard	-0.03	76	0.01	0.00	-0.77

F. SURVEY RESPONSES

As discussed in Chapter III, each study participant completed a written survey after the conclusion of the experiment. The survey consisted of twelve questions, requiring responses in short-answer, narrative form, and on a 1 – 7 Likert Scale (1 = lowest/least positive response, 7 = greatest/most positive response). The survey questions were focused on service information and experience, sleep habits, and demographic information (addressed in Chapter III.A.2). As before, for responses made on the 1 – 7 Likert Scale, the reciprocal (1 = 7, 2 = 6, etc.) and $LOG_{10}(x)$ of the resulting figure was used for inputs into the ANOVA. However, raw Likert Scale figures were used for the descriptive statistics. Furthermore, any extreme outliers that were identified after transformation were omitted.

1. Efficiency of TMPLAR Application

The first question asked of the participants in the survey was, "How efficient was the planning software?" As this question required a response on a Likert Scale, the aforementioned transformation was applied to the responses. A univariate ANOVA conducted the transformed responses indicated no significant main effect of efficiency ratings across DSS level; $F(2, 37) = 0.83$, $p = 0.44$, and $\eta_p^2 = 0.04$. See Table 14 for descriptive statistics and pairwise comparison of user ratings of efficiency of TMPLAR application.

Table 14. User Efficiency Ratings of TMPLAR Application.

User Mean TMPLAR Efficiency Ratings

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	14	5.86	1.10	0.29	5.22	6.49
DSS 1	13	5.54	0.78	0.22	5.07	6.01
DSS 3	13	5.46	1.27	0.35	4.70	6.23

2. Helpfulness of TMPLAR Application

The second question on the post-experiment survey was, "Do you feel that the software was helpful in enabling you to choose the optimum route?" As this question required a response on a Likert Scale, the descriptive statistics are expressed in terms of raw figures while the Univariate ANVOA utilizes the aforementioned transformed data set. Again, no significant main effect of helpfulness ratings across DSS level; $F(2, 31) = 0.17$, $p = 0.85$, and $\eta_p^2 = 0.01$, was discovered. As a result, DSS can be said to have no effect on user helpfulness ratings. See Table 15 for descriptive statistics of TMPLAR application helpfulness.

Table 15. User Helpfulness Ratings of TMPLAR Application.

User Mean TMPLAR Helpfulness Ratings

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	8	6.00	0.00	0.00	6.00	6.00
DSS 1	13	6.00	1.00	0.28	5.40	6.60
DSS 3	13	6.17	0.72	0.21	5.71	6.62

3. Time Available to Complete Scenarios

The next question contained in the participant survey asked the users, "Did you feel like you had enough time to plan all the routes?" Although users were briefed to take as much time as needed, they were required to provide a response on a Likert Scale. A Univariate ANOVA reported no significant main effect of efficiency ratings across DSS

level; $F(2, 37) = 0.89$, $p = 0.42$, and $\eta_p^2 = 0.05$. Therefore, DSS was found to have no effect on the users' perception of time available to complete the assigned task. Refer to Table 16 for descriptive statistics of user ratings of perceived time available to complete the assigned task.

Table 16. User Ratings of Time Available to Complete Scenarios.

User Ratings of Time Available to Complete Scenarios

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	14	6.07	1.07	0.29	5.45	6.69
DSS 1	13	6.38	0.65	0.18	5.99	6.78
DSS 3	13	6.54	0.88	0.24	6.01	7.07

4. User-Proposed Additions to TMPLAR Functionality

The first question on the post-experiment survey that required a narrative answer was, "Are there any additions to the software that would have helped you in your route decisions?" Of the 40 participants that completed the post-experiment survey, 21 provided substantive responses. The received responses generally addressed three aspects of TMPLAR: the user interface, the map, and the presentation of data.

User suggestions addressing the user interface were centered on the ability to customize the screen or specific fields to suit individual taste. A common theme was the ability to rearrange the order of the data columns, or completely hide them, within the table field to more suit the individual planner's mental priorities. Additionally, a single button to display or hide all routes on the map and distribution function field were proposed. For the DSS 3 interface specifically, it was proposed that a simpler way to select routes on the Pareto distribution should be implemented, such as by drawing a selection box with the mouse over the desired points. Also, the ability to sort routes by two different fields (maximum wave height and the by fuel usage, for example) was suggested. Finally, the ability to set the Pareto distribution and the cumulative distribution function to display maximum simulation values instead of medians was suggested to allow the user to better conceptualize the overall risk of selected routes.

On the other hand, user suggestions pertaining to the map field were comparatively simple. Multiple users expressed the desire for weather affecting navigation routes to be visually displayed on the map. Other requests included the incorporation of the surface tracks of known contacts and a visual depiction of traffic density.

The final and most substantive classification of user suggestions for the improvement of the TMPLAR application relates to the presentation of data. The most recurring comment was that ship speed was not displayed anywhere on the screen. This omission should be corrected as speed is an important consideration for safety of transit. Another suggestion that was received on more than one occasion was for the display of the AI's relative confidence in the displayed route's vitals. If the route was calculated using unreliable or incomplete data, the users generally need to know before making decisions. Another suggestion was for an option to set the most fuel efficient or the more direct route as the system default, then display all other calculated route statistics in terms of the default route. For example, "Route 1," would always be set as the default. "Route 2," would then display distance to destination as, "+150 nm," and so on. This would allow for easier mental comparison of multiple route options. Finally, multiple users requested that routes outside of a certain confidence interval be omitted to reduce metal load on the decision maker as they weigh their options.

5. User Sleep Hours Prior to Experiment

The last item on the survey addressed the amount of sleep the users experienced the night prior. No transformations were required for the user sleep hours data set. The Univariate ANOVA found no evidence of a significant main effect in reported user sleep hours across DSS level; $F(2, 37) = 0.97$, $p = 0.39$, and $\eta_p^2 = 0.05$. Therefore, it was determined that a significant difference is sleep hour prior to the experiment could not be found across DSS users. As a result, hours of sleep prior to the experiment cannot be examined as a contributing factor for time to decision performance or cognitive load. Refer to Table 17 for descriptive statistics of user sleep hours prior to experiment across DSS levels.

Table 17. User Hours of Sleep Prior to Experiment across DSS Levels.

User Hours of Sleep Prior to Experiment Across DSS

Descriptive Statistics						
Field	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval	
					Lower Bound	Upper Bound
DSS 0	14	7.00	1.09	0.29	6.37	7.63
DSS 1	13	6.42	1.32	0.37	5.63	7.22
DSS 3	13	6.96	1.16	0.32	6.26	7.66

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

1. Time to Decision

The built-in functionality of the TMPLAR application and the Gazepoint software suite allowed for many conclusions to be drawn from the experimental trials. The first and most important of these conclusions pertains to the time to decision achieved by users of AI-augmented and baseline DSS. The results show that DSS 1 and DSS 0 users recorded the fastest time to decision with DSS 3 users posting the slowest times. However, users of the baseline system (DSS 0) and the simplified, AI-Augmented DSS (DSS 1) did not show a significant difference in time to decision. This indicates that DSS 0 users were provided with enough information to facilitate quick decision-making while DSS 1 users reaped the benefits of the AI-augmentation without the drawbacks of the increased user interface complexity of DSS 3. The result also indicates that simpler displays, with or without AI-augmentation, are more conducive to faster time to decision than more complex, information-dense interfaces.

Easy (projected fuel usage within one standard deviation of the mean for the given route), Medium (projected fuel usage within two standard deviations of the mean), and Hard (all COAs across the entire distribution) difficulties were also assessed for their effects on time to decision. Significant differences in time to decision were shown to exist between the Easy and Hard time to decision data points. This indicates better, in terms of time, decision-making capability when the DSS restricts the route options to those fitting the route output parameters (routes within one standard deviation of the mean) as opposed to displaying the entire solution set as calculated by TMPLAR. This clear result represents a future TMPLAR setting that can default route options output to those within default setting of one standard deviation from the calculated mean with a setting that can be adjusted to display the entire distribution if desired.

2. User Interaction with Screen Fields

User interaction with the TMPLAR interface was quantified by recording mouse clicks and visual fixations on or within pre-defined screen fields for all DSS levels. For DSS 0, users clicked the computer mouse within the dimensions of the Table field at a frequency that was greater to a statistically significant degree than the Map Field. This data is reinforced by similar results for visual fixations, which indicated that users fixated on the Table at a statistically significant higher rate than the Map field. From both the click data and the eye tracking information, it is concluded that users interacted with and relied upon the data table most in their decision making with DSS 0.

A similar result was seen in DSS 1, where the screen was divided into three fields, the Map, a Cumulative Distribution Function, and a Table field. User click data indicated that experiment participants clicked the mouse within the Table field at higher, statistically significant rates than either the Map (which exhibited the second greatest click rates) or the Cumulative Distribution Function (lowest click rates) fields. Furthermore, a significant difference existed between the Map and Cumulative Distribution Function fields, with nearly no user clicks in the latter. Visual fixation data obtained from the eye tracker again reinforces the click data by ordering the level of user fixations (from greatest to least) as the Table, Map, and Cumulative Distribution Function fields. In DSS 1 visual fixations, statistically significant differences existed between the Map/Table and Cumulative Distribution Function/Table Fields. It can therefore be concluded that in DSS 1, users physically and visually interacted with the Table field at statistically significant higher rates than the Map or Cumulative Distribution Function field.

Results for DSS 3 user interaction with screen fields again demonstrate the primacy of the Table field for decision makers. Click data showed that users interacted with the interface in the following descending order of frequency: Table, Pareto Distribution. Map and finally the Cumulative Distribution function. Statistically significant differences were present for user clicks between all screen fields except the Map and the CDF. DSS 3 user visual fixations yielded similar results with gaze rates for fields being recorded in the following descending order: Table, Pareto Distribution, Table, Pareto Distribution, Cumulative Distribution Function, and lastly Map. A statistically significant difference

existed between all screen fields except the Map and CDF, as in the user click analysis. From these results, it is concluded that DSS 3 users physically and visually interacted with the Table and Pareto Distribution at statistically significant higher rates than the Map or Cumulative Distribution Function field. However, it is important to note, that users were required to click on route instances on the Pareto Distribution to progress through the scenarios. As a result, data skewness toward the Pareto Distribution is artificial and accounted for by the DSS 3 user interface design.

An analysis of user interaction with screen fields that were common to all three DSS levels (Map and Table fields) yielded results indicative of those found in the within-DSS examinations. Both user clicks and gaze fixations favored the Table field at statistically significant and higher rates than those recorded in the Map Field. This result and those above indicate that users both tactilely through an input device and physiologically through eye movements favor data presented in a tabulated, numeric format when presented with tasks that require a comparative analysis of mutually exclusive, multi-attribute options.

3. Cognitive Load

Pupil diameter measurements, which are an indicator of cognitive load, were measured with the Gazepoint suite for users across all DSS levels. Mean pupil diameter indicated that DSS 3 users were under the greatest cognitive load, followed by DSS 0 users, and DSS 1 users, who experienced the least cognitive load. Statistically significant differences existed between users of all DSS levels and the cognitive load results are congruent with those of the time to decision analysis (DSS 3: longest time required/greatest cognitive load; DSS 0: moderate time required/moderate cognitive load, DSS 1: shortest time require/least cognitive load).

4. User Confidence Ratings and Survey Information

User confidence ratings exhibited statistically significant differences between those recorded in the three DSS levels. DSS 0 received the highest confidence ratings for route selections, followed by DSS 3, with DSS 1 receiving the lowest confidence ratings. It is important to note that these user ratings do not correspond with the ordinal rankings of time

to decision performance or cognitive load assessments made above. Confidence ratings across scenario difficulty were determined to be highest for Easy difficulty, next highest for Medium difficulty, and lowest for Hard difficulty. The differences calculated between all three difficulty ratings was shown to be statistically significant. As a result, it was concluded that users exhibit greater confidence in the selections made, aided by a DSS, when possible solutions are restricted to within one or two standard deviations of the mean as opposed to an unrestricted solution set. Also, user ratings of the TMPLAR application's efficiency and helpfulness did not differ significantly across DSS level. Finally, User self-assessments of their own maritime route planning experience were significantly higher for DSS 1 than DSS 0 or DSS 3. However, because, "correctness" of route selection was not taken into account in the analysis, the greater relative level of experience did not impact the results of what was analyzed (time to decision, cognitive load, reliance on data fields).

5. Overall Assessment

Given the results of the experiment and the ensuing analysis, it was determined that the DSS 1 interface produced the most beneficial results as compared to the alternatives. The underlying function of DSS 1 was augmented with AI that ran maritime route simulations against known and forecasted METOC conditions, resulting in options that were more refined and relevant to the simulated operational environment. DSS 0 did not provide this functionality at all, and while DSS 3 did possess the same underlying AI-algorithm, its more complex user interface and presentation resulted in poorer performance during experimental trials. As a result, DSS 1 users exhibited faster time to decision than the alternatives, significantly so in the case of DSS 3. Additionally, DSS 1 had shown by a statistically significant margin that its user interface produced a level of cognitive load that was significantly less than DSS 0 and DSS 3.

Nested within the examination pertaining to DSS levels, specific data fields were also examined for the level of interaction users had with each. In every DSS level, it was shown that users relied most heavily and by a statistically significant margin on route data in a standard table format. Where available (as in DSS 3), route data displayed in Pareto Distribution was also heavily utilized by users, but this result can be explained by the

requirement of users to select routes on the Pareto Distribution to initiate route comparison in each scenario.

B. POTENTIAL BENEFITS

In addition to the benefits of greater operational mobility afforded by minimized fuel consumption and more efficient routing, this study revealed several additional significant benefits. From the time to decision analysis, it was shown that AI-augmentation to DSS has the ability to significantly reduce the time it takes for a user to arrive at a decision over a baseline application. This can focus the development and acquisition of future DSS in the DON, thereby maximizing performance out of human users while saving time and real-world resources. Also, the inclusion of capabilities such as the tracks of surface contacts, a real-time moving map, and threat range rings will make the TMPLAR application not only an effective mission planning tool, but also an effective common operational picture application. Additionally, the balance demonstrated between user interface complexity and volume of data displayed demonstrated an ideal compromise in which users were most readily able to make calculated decisions while experiencing the least cognitive load, thereby allowing the decision-maker more time to dedicate to more dynamic tasks.

A major benefit yet to be realized from a DSS like TMPLAR is its applications outside of the surface maritime domain. Aside from its obvious uses for subsurface vessels, TMPLAR, or a similar application, can be conceivably adapted for use in land environments. The ability to calculate fuel requirements for surface routes while taking into consideration inputs such as METOC effects, road conditions, threat posture, visibility, etc., would be in high demand for any operational environment where multi-vehicle mounted movements are routinely conducted.

C. LIMITATIONS

There are several shortcomings of this study that stem from the design of the version of the TMPLAR application used for the experiment. A major shortcoming with the TMPLAR application that limited a result of the experiment was that users were required to interact with certain screen fields to progress the scenario, specifically, the

Pareto Distribution in DSS 3. In that iteration of TMPLAR, users were required to select a route on the distribution to make its vitals appear in the Table field and its track on the Map field. This caused artificially inflated values for the level of interaction both through mouse clicks and visual fixations for that screen field. As a result, this study does not adequately address the true relative value of a Pareto Distribution compared to a data table, map, or cumulative distribution function in navigation problems. Another shortcoming is the inclusion of only five origin/destination pairs for navigation scenarios. This lack of variety led to feelings of monotony in the participants as they completed the required 60 scenarios. If more five origin/destination pairs were included, the participants would have been continuously presented with fresh, challenging navigation problems. This in turn could have limited the inclination of participants to complete the scenarios faster to finish the experiment trial sooner. The final limitation of this study was the lack of any practical way to assess or, "score," the experiment participants' route selections. If there was a feasible way to determine is a user selected the correct route, or even better, a route scored on a spectrum of correctness given the known conditions, the researcher would have had a powerful metric to assess the relative effectiveness of the TMPLAR DSS levels.

D. RECOMMENDATIONS & FUTURE WORK

It is highly recommended that future NPS student researchers continue studies into human-machine teaming and AI-augmentation to DSS in partnership with NJIT, UCONN, and NRL-MRY. The results of this study are compelling and show definite promise for moving the body of knowledge forward in the field of military applications of AI in operational environments. The TMPLAR application, although useful in its current state, has much room for improvement in the greater levels of performance it can provide to the user. The largest recommendation for further study is to quantify the relative level of performance gained, in terms of correct route choices, by employing an AI-augmentation to a DSS over a standard, non-augmented system. Additionally, it is advised that future studies allow for the option of user customization of the interface prior to each trial as a separate measure in addition to set levels of DSS. Also, biometric measurements such as heart rate and Cortisol can be measured for participants before, during, and after experiment trails to gain a greater understanding of stress levels during the experiment and

how they affect performance and cognitive load of the users. Lastly, it is recommended that in the future, each experiment participant completes a set of fewer scenarios at each DSS level, to include the customized option, with the goal of achieving a greater number of participants per DSS.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. EXIT QUESTIONNAIRE

Exit Questionnaire

User ID: _____

1. How efficient was the planning software? (Circle one choice)

   Least Efficient 1 2 3 4 5 6 7 Most Efficient

2. Do you feel that the software was helpful in enabling you to choose the optimum route? (Circle one choice)

   Least Helpful 1 2 3 4 5 6 7 Most Helpful

3. Did you feel like you had enough time to plan all the routes? (Circle one choice)

   Too Little Time 1 2 3 4 5 6 7 Sufficient Time

4. Are there any additions to the software that would have helped you in your route decisions?

5. What is your age?

6. What is your gender?

7. What branch of service? (If any)

8. What is your experience level in maritime route planning?

   No Experience 1 2 3 4 5 6 7 Very Experienced

   a. If you have done maritime route planning in the past, did you use software or manual planning? Please name the software application if applicable.

9. How many hours of sleep did you get last night?

   a. Is this more/less/same as usual?

10. What time did you wake up today?

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. INFORMED CONSENT FORM

Naval Postgraduate School
Consent to Participate in Research

Introduction. You are invited to participate in a research study entitled Human-AI Symbiosis. The purpose of the research is to evaluate the effectiveness of an AI enhanced decision support system for maritime planning.

1) Participation is voluntary. Refusal to participate will involve no penalty or loss of benefits to which you would otherwise be entitled, and you may discontinue participation at any time without penalty or loss of benefits to which you otherwise would be entitled.
2) The purpose of this study is to evaluate an enhanced decision making planning tool.
3) There are no foreseeable risks or discomforts to you as a participant in this study.
4) You will receive no direct benefits from participating in this study, but you will learn about maritime planning systems.

Procedure. The study will require approximately an hour a half of your time in which you will learn maritime planning software, and be asked to plan 60 routes will varying degrees of uncertainty. Half of the participants will use the planning software with enhanced AI decision support, and half of the participants will use the planning software as is. Approximately 60 participants will be needed for the study.

Location. The interview/survey/experiment will take place at the Naval Postgraduate School, Glasgow 103.

Cost. There is no cost to participate in this research study.

Compensation for Participation. No tangible compensation will be given.

Confidentiality & Privacy Act. Any information that is obtained during this study will be kept confidential to the full extent permitted by law. All efforts, within reason, will be made to keep your personal information in your research record confidential but total confidentiality cannot be guaranteed. The only information that will have your name on it is the informed consent, and that will not be able to be matched with your data. Your data will be anonymous.

Points of Contact. If you have any questions or comments about the research, or you experience an injury or have questions about any discomforts that you experience while taking part in this study please contact the Principal Investigator, Dr. Mollie McGuire, mrmcguir@nps.edu. Questions about your rights as a research subject or any other concerns may be addressed to the Navy Postgraduate School IRB Chair, Dr. Larry Shattuck, 831-656-2473, lgshattu@nps.edu.

Statement of Consent. I have read the information provided above. I have been given the opportunity to ask questions and all the questions have been answered to my satisfaction. I have been provided a copy of this form for my records and I agree to participate in this study. I understand that by agreeing to participate in this research and signing this form, I do not waive any of my legal rights.

_____ _____
Signature of Participant Date

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. TMPLAR EXPERIMENT VERBAL BRIEF

Tool for Multi-Objective Planning and Asset Routing (TMPLAR) Experiment Instructions

Purpose: The purpose of this study is to evaluate the effectiveness of UCONN/NJIT-developed, AI-augmentation to the TMPLAR application for surface vessel route planning/deconfliction. Experiment participants will validate the efficacy of the AI-augmented TMPLAR application as compared to the baseline TMPLAR Decision Support System (DSS).

Experimental Procedure: As a participant, you will be assigned one of three levels of DSS interface for the experiment (user guide provided in binder). Upon receiving your User ID from the administrator, enter the assigned User ID into the TMPLAR home screen. The TMPLAR application will then navigate to the respective experimental interface that represents the level of DSS that you have been assigned.

Users will select the routes that they deem best fulfill the commander's guidance (given below) in 60 total route planning scenarios. The scenarios will vary according to vessel point of origin/destination and relative level of Meteorological and Oceanographic (METOC) uncertainty along the given route. Users will not be limited on time available to select a route for a given scenario. After a route is chosen, you will be prompted to enter relative level of confidence in the appropriateness of your selection based off the known/forecasted METOC parameters.

Task: For each scenario, compare the DSS-generated surface routes and consider the relevant attributes that will affect your ship's ability to arrive at its destination at the earliest time while expending the least possible amount of fuel.

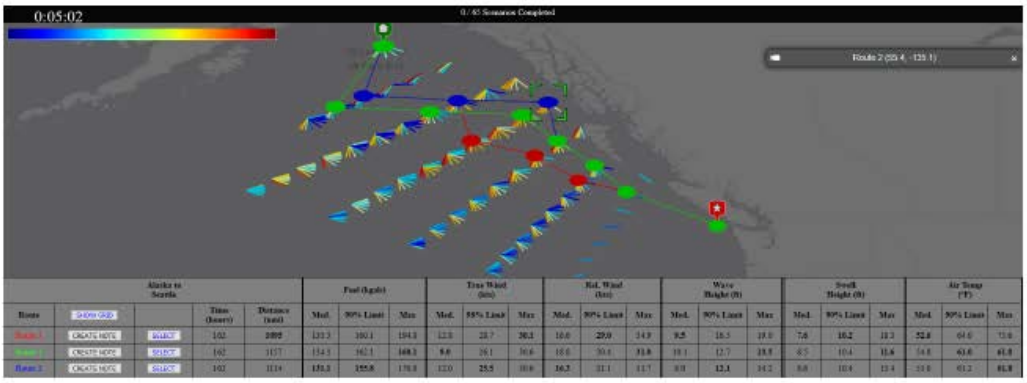
- **Constraints/Restraints and Assumptions:**
  - **Users are to select the route that they believe is the optimal balance between fuel expenditure and distance**
  - Ship cannot transit through conditions that exceed thresholds for safe transit
    - True Wind: 35 knots
    - Waves: 39 feet
  - Internal factors effecting fuel efficiency are optimized in the simulation; the simulated ship's powerplant is configured for most efficient travel given the known and forecasted METOC conditions and hull fouling is not present on the simulated vessel
  - Commercial vessel traffic in confined waterways and potential enemy threats are not considered in the simulation

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. TMPLAR (DSS 0) USER MANUAL

TMPLAR Manual

Developed by: The University of Connecticut (UConn)



Alaska to Seattle					Fuel (kgals)			True Wind (kts)			Rel. Wind (kts)			Wave Height (ft)			Swell Height (ft)			Air Temp (°F)		
Route	SHOW/HIDE		Time (hours)	Distance (nmi)	Med.	80% Limit	Max	Med.	80% Limit	Max	Med.	80% Limit	Max	Med.	80% Limit	Max	Med.	80% Limit	Max	Med.	80% Limit	Max
Route 1	CREATE/HIDE	SELECT	102	1095	133.3	160.1	194.8	12.8	20.7	30.1	10.0	20.0	34.9	8.5	16.5	19.0	7.6	10.2	12.3	52.8	64.0	73.0
Route 2	CREATE/HIDE	SELECT	102	1157	134.5	162.1	168.3	9.0	19.1	30.6	10.0	20.6	31.0	10.1	12.7	15.5	8.5	10.4	11.6	54.0	63.0	69.8
Route 3	CREATE/HIDE	SELECT	102	1114	131.3	155.0	178.8	12.0	25.5	30.0	16.2	22.1	32.7	8.0	12.1	14.2	8.0	10.4	13.4	53.0	63.2	81.8

Last updated: 9-25-2019 by S. Uziel

Intro Screen

At the beginning of the experiment, the above screen will be displayed. Enter your given ID number and click start.

Route Info Table

Alaska to Seattle					Fuel (kgals)			True Wind (kts)			Rel. Wind (kts)			Wave Height (ft)			Swell Height (ft)			Air Temp (°F)		
Route	SHOW QMD		Time (hours)	Distance (nmi)	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max	Med.	90% Limit	Max
Route 1	CREATE ROUTE	SELECT	162	**1095**	155.3	160.1	194.0	12.8	28.7	**30.1**	16.6	**29.0**	34.9	**9.5**	16.5	19.0	7.6	**10.2**	18.3	**52.6**	64.0	73.6
Route 3	CREATE ROUTE	SELECT	162	1157	154.1	162.1	**168.1**	**9.0**	28.1	30.6	18.8	30.4	**31.8**	10.1	12.7	**13.5**	8.5	10.4	**11.6**	54.8	**61.0**	**61.0**
Route 2	CREATE ROUTE	SELECT	162	1314	**151.1**	**155.8**	170.0	12.0	**25.5**	30.6	**16.3**	31.1	31.7	9.9	**12.1**	14.2	8.6	10.4	13.4	55.0	61.2	**61.8**

The route info table displays the following information about the routes. Double-clicking on any of the fields will allow the user to sort the data in ascending or descending order. Bold values within fields indicate the most favorable value for routes within a given attribute.

- **Time of Passage**: Total time in hours that transit is calculated to require
- **Distance**: Total distance in nautical miles from the start point to the destination
- **Fuel**: Total fuel cost in thousands of gallons (median, 90th percentile, and maximum values given for all simulations)
- **True Wind Speed**: Wind Speed in knots. Wind speed has a maximum allowable value of **35 knots** for safe transit.
- **Relative Wind Speed**: Wind Speed in knots relative to the speed and direction of travel.
- **Wave Height**: Significant wave height in feet (the mean wave height of the highest third of the waves). Significant wave height has a maximum allowable value of **39 feet** for safe transit.
- **Swell Height**: Primary swell height in feet.
- **Air Temp**: Air temperature in Fahrenheit

All values in the table display the following three measurements:

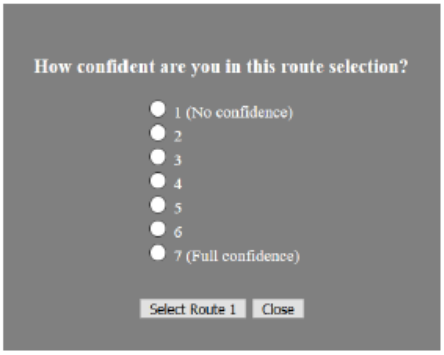
- **Median**: Value lying at the midpoint of the frequency distribution within the simulation.
- **90th Percentile**: The given value that ranks greater than 90% of the simulated data within a specific field.
- **Max**: Maximum value for a given attribute at any point in the route for any possible weather forecast.

In addition, the route info table has the following controls:

- **Route #**: Clicking the button that displays the route number toggles display of the route on the map.
- **Show Grid:** Toggles on/off route vectors. Route vectors indicate favorability from various waypoint approach angles due to winds and current (red = least favorable, blue = most favorable)

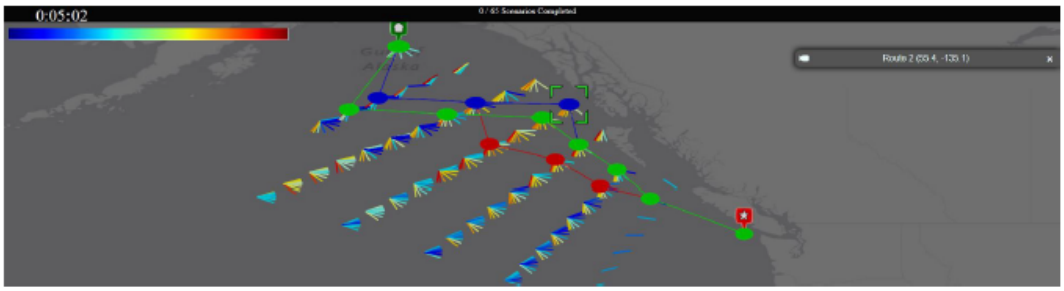
- **Create Note**: Allows the user to scribe notes for personal reference and route comparison (shown below).

- **Select**: Selects the route as your choice for the route for the current scenario. This leads to a confirmation prompt where you rate your confidence in the route that you have selected between 1 (no confidence)and 7 (full confidence), as shown below.



Map Display

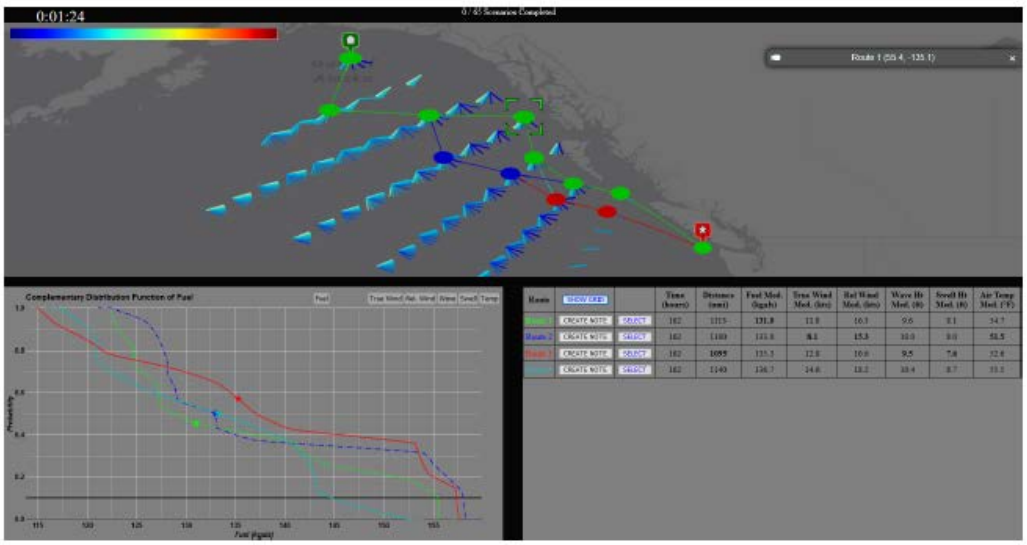
The map displays the waypoints for the routes that have been added from the Route Info Table. Left clicking and dragging will pan the view of the map. To zoom the map, right click and drag, or use the scroll wheel. Double-clicking on a route waypoint will allow user to clarify which routes are designated on the map and display latitude/longitude locations of a selected waypoint.



APPENDIX E. TMPLAR (DSS 1) USER MANUAL

TMPLAR Manual

Developed by: The University of Connecticut (UConn)



Last updated: 9-25-2019 by S. Uziel

Intro Screen

At the beginning of the experiment, the above screen will be displayed. Enter your given ID number and click start.

Route Info Table

Route	SHOW GRID		Time (hours)	Distance (nmi)	Fuel Med. (kgals)	True Wind Med. (kts)	Rel Wind Med. (kts)	Wave Ht Med. (ft)	Swell Ht Med. (ft)	Air Temp Med. (°F)
Route 1	CREATE NOTE	SELECT	162	1113	**131.0**	11.8	16.3	9.6	8.1	54.7
Route 2	CREATE NOTE	SELECT	162	1100	133.0	**8.1**	**15.3**	10.0	8.0	**51.5**
Route 3	CREATE NOTE	SELECT	162	**1095**	135.3	12.8	16.6	**9.5**	**7.6**	52.6
Route 4	CREATE NOTE	SELECT	162	1140	136.7	14.6	18.2	10.4	8.7	55.5

The route info table displays the following information about the routes. Double-clicking on any of the fields will allow the user to sort the data in ascending or descending order. Bold values within fields indicate the most favorable value for routes within a given attribute.

- **Time of Passage**: Total time in hours that transit is calculated to require
- **Distance**: Total distance in nautical miles from the start point to the destination
- **Fuel**: Total fuel cost in thousands of gallons (median)
- **True Wind Speed**: Wind Speed in knots. Wind speed has a maximum allowable value of **35 knots** for safe transit.
- **Relative Wind Speed**: Wind Speed in knots relative to the speed and direction of travel.
- **Wave Height**: Significant wave height in feet (the mean wave height of the highest third of the waves). Significant wave height has a maximum allowable value of **39 feet** for safe transit.
- **Swell Height**: Primary swell height in feet.
- **Air Temp**: Air temperature in Fahrenheit

All values in the table display the following measurements:

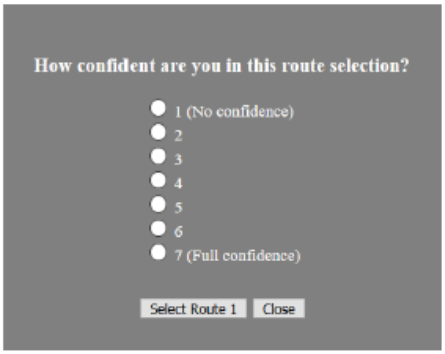
- **Median**: Value lying at the midpoint of the frequency distribution within the simulation.

In addition, the route info table has the following controls:

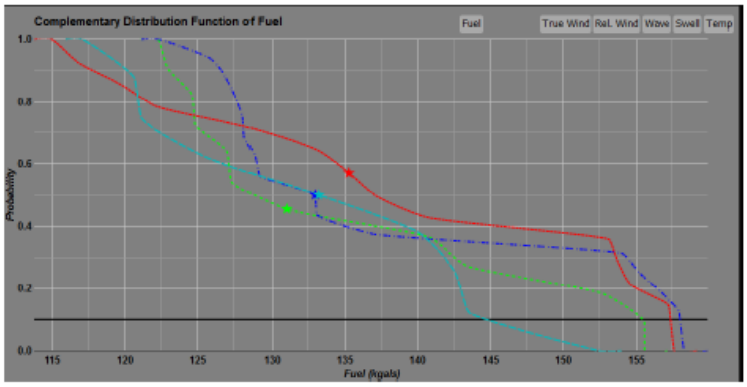
- **Route #**: Clicking the button that displays the route number toggles display of the route on the map.
- **Show Grid:** Toggles on/off route vectors. Route vectors indicate favorability from various waypoint approach angles due to winds and current (red = least favorable, blue = most favorable).

- **Create Note**: Allows the user to scribe notes for personal reference and route comparison (shown below).

- **Select**: Selects the route as your choice for the route for the current scenario. This leads to a confirmation prompt where you rate your confidence in the route that you have selected between 1 (no confidence) and 7 (full confidence), as shown below.



Cumulative Distribution Function (CDF) Graph



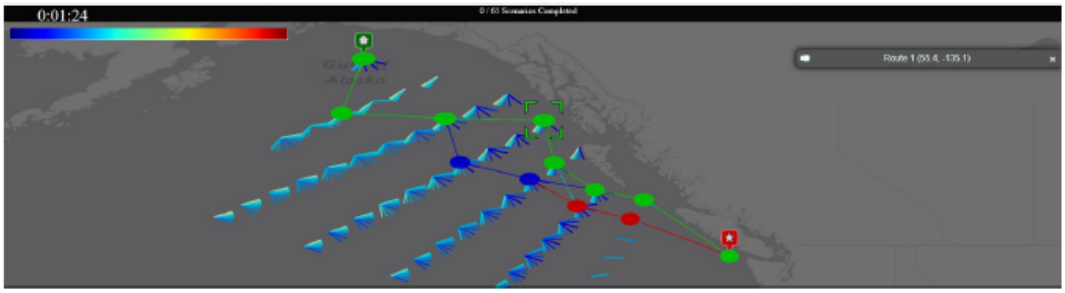
The Cumulative Distribution Function Graph is a display that represents the probability distribution of values that change over the weather forecasts. Each point in the line represents the probability that the attribute value will be more than the value at that point. For example, in the image above, the red line shows that there is a 70% (0.7) probability that the fuel will be less than 130 kgals for the route represented by the red line. Stars on the lines indicate median values of computer simulations for a given route.

If a line is fully to the right of another line, this means that the line to the right has a greater probability of having higher fuel cost than the left line. If the lines cross, this means that the same route does not always have a higher probability of having a lower fuel cost (for a given level of fuel expenditure).

Hovering over a line on the graph will show the specific probability numbers of the point as well as the route number. At the top right of the graph are tabs to select which value the CDF represents. The values are the same as the values in the Route Info Table.

Map Display

The map displays the waypoints for the routes that have been added from the Route Info Table. Left clicking and dragging will pan the view of the map. To zoom the map, right click and drag, or use the scroll wheel. Double-clicking on a route waypoint will allow user to clarify which routes are designated on the map and display latitude/longitude locations of a selected waypoint.

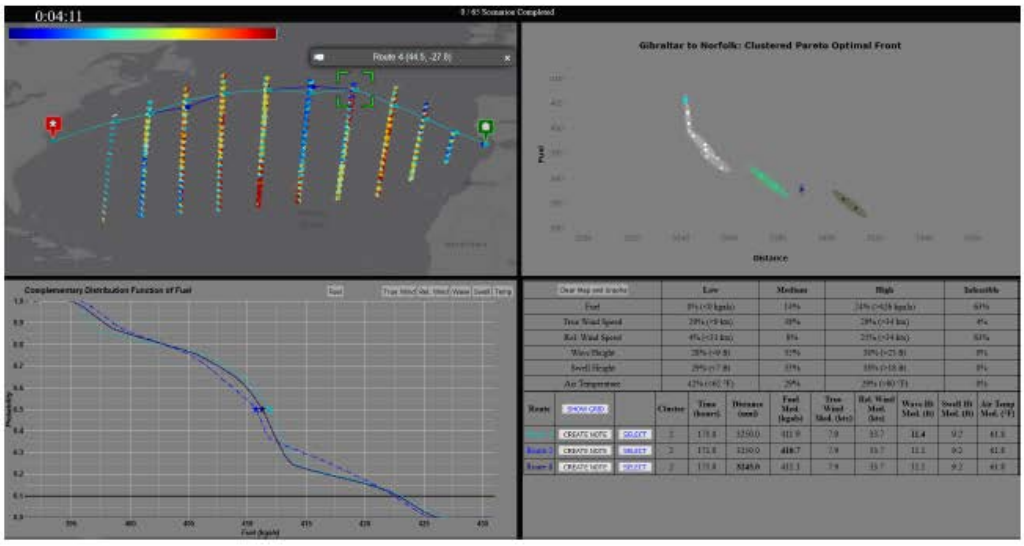


THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX F. TMPLAR (DSS 3) USER MANUAL

TMPLAR Manual

Developed by: The University of Connecticut (UConn)

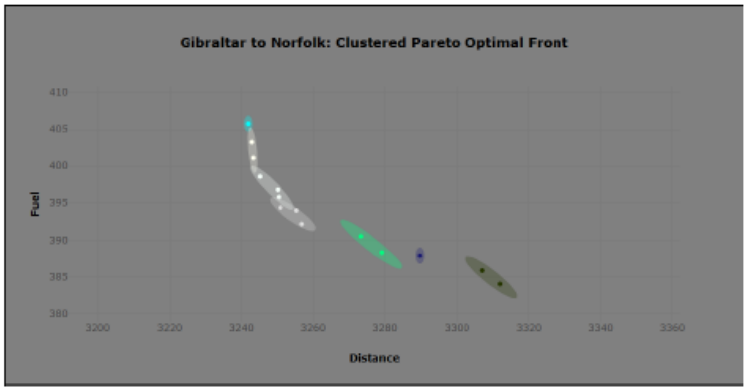


Last updated: 9-25-2019 by S. Uziel

Intro Screen

At the beginning of the experiment, the above screen will be displayed. Enter your given ID number and click start.

Pareto Optimal Front



This display shows fuel usage compared to distance of all possible routes, with ranges of how much the fuel may vary over the different possible weather conditions. Clicking on the points in one of the clusters adds the routes in that cluster to the Route Info Table.

Route Info Table

Clear Map and Graphs			Low	Medium	High	Infeasible
Fuel			0% (<0 kgals)	14%	24% (>426 kgals)	63%
True Wind Speed			29% (<9 kts)	38%	29% (>34 kts)	4%
Rel. Wind Speed			4% (<31 kts)	8%	25% (>34 kts)	63%
Wave Height			28% (<9 ft)	35%	38% (>23 ft)	0%
Swell Height			29% (<7 ft)	33%	38% (>18 ft)	0%
Air Temperature			42% (<65 °F)	29%	29% (>80 °F)	0%

Route	SHOW GRID		Cluster	Time (hours)	Distance (nmi)	Fuel Med. (kgals)	True Wind Med. (kts)	Rel. Wind Med. (kts)	Wave Ht Med. (ft)	Swell Ht Med. (ft)	Air Temp Med. (°F)
Route 1	CREATE NOTE	SELECT	2	171.0	3250.0	411.9	7.9	33.7	**11.4**	9.2	61.8
Route 2	CREATE NOTE	SELECT	2	171.0	3250.0	**410.7**	7.9	33.7	11.5	9.2	61.8
Route 8	CREATE NOTE	SELECT	2	171.0	**3245.0**	411.3	7.9	33.7	11.5	9.2	61.8

The route info table displays the following information about the routes. Attributes are categorized as "Low", "Medium", and "High" based on the distribution of the attributes across all possible routes. Double-clicking on any of the fields will allow the user to sort the data in ascending or descending order. Bold values within fields indicate the most favorable value for routes within a given attribute.

- **Time of Passage**: Total time in hours that transit is calculated to require
- **Distance**: Total distance in nautical miles from the start point to the destination
- **Fuel**: Total fuel cost in thousands of gallons
- **True Wind Speed**: Wind Speed in knots. Wind speed has a maximum allowable value of **35 knots** for safe transit.
- **Relative Wind Speed**: Wind Speed in knots relative to the speed and direction of travel.
- **Wave Height**: Significant wave height in feet (the mean wave height of the highest third of the waves). Significant wave height has a maximum allowable value of **39 feet** for safe transit.
- **Swell Height**: Primary swell height in feet.
- **Air Temp**: Air temperature in Fahrenheit

All values in the table display the following measurements:

- **Median**: Value lying at the midpoint of the frequency distribution within the simulation.

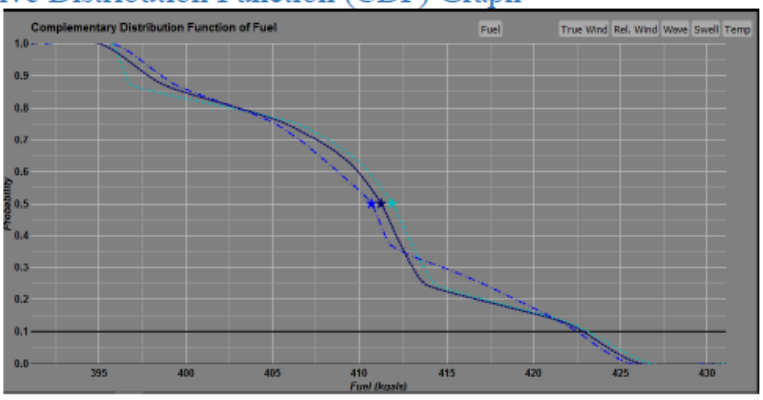
In addition, the route info table has the following controls:

- **Route #**: Clicking the button that displays the route number toggles display of the route on the map.
- **Show Grid:** Toggles on/off route vectors. Route vectors indicate favorability from various waypoint approach angles due to winds and current (red = least favorable, blue = most favorable).

- **Create Note**: Allows the user to scribe notes for personal reference and route comparison (shown below).

- **Select**: Selects the route as your choice for the route for the current scenario. This leads to a confirmation prompt where you rate your confidence in the route that you have selected between 1 (no confidence)and 7 (full confidence), as shown below.

Cumulative Distribution Function (CDF) Graph



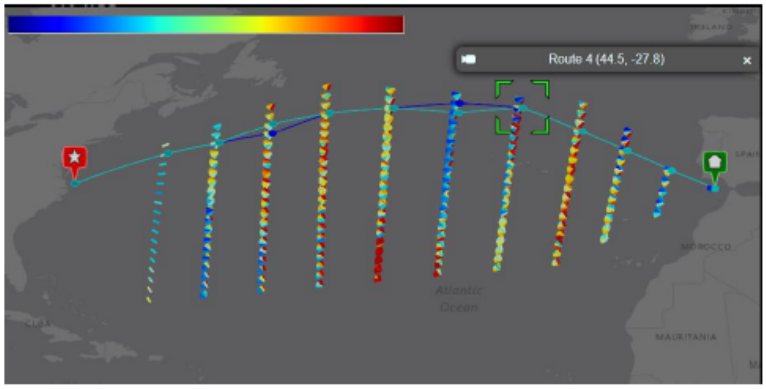
The Cumulative Distribution Function Graph is a display that represents the probability distribution of values that change over the weather forecasts. Each point in the line represents the probability that the attribute value will be more than the value at that point. For example, in the image above, the cyan line shows that there is a 63% (0.63) probability that the fuel will be less than 410 kgals for the route represented by the cyan line. Stars on the lines indicate median values of computer simulations for a given route.

If a line is fully to the right of another line, this means that the line to the right has a greater probability of having higher fuel cost than the left line. If the lines cross, this means that the same route does not always have a higher probability of having a lower fuel cost (for a given level of fuel expenditure).

Hovering over a line on the graph will show the specific probability numbers of the point as well as the route number. At the top right of the graph are tabs to select which value the CDF represents. The values are the same as the values in the Route Info Table.

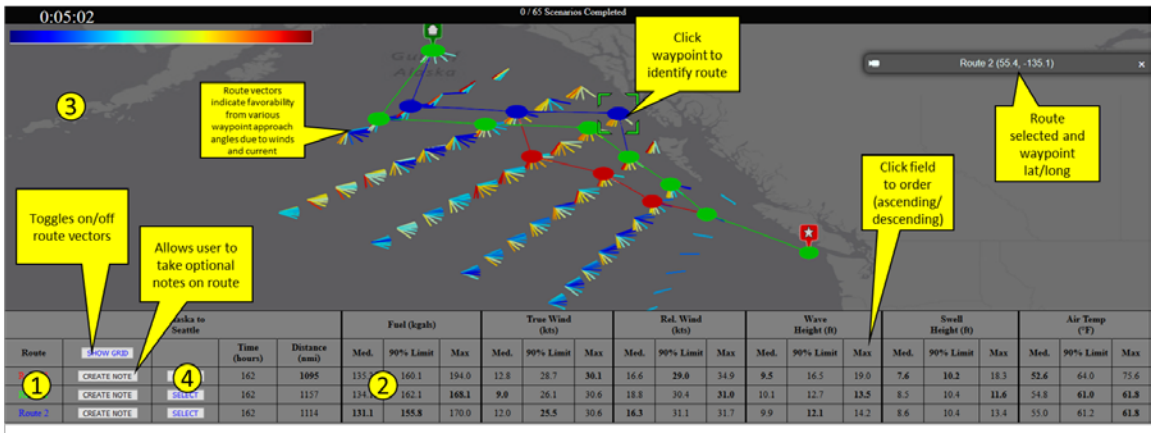
Map Display

The map displays the waypoints for the routes that have been added from the Route Info Table. Left clicking and dragging will pan the view of the map. To zoom the map, right click and drag, or use the scroll wheel. Double-clicking on a route waypoint will allow user to clarify which routes are designated on the map and display latitude/longitude locations of a selected waypoint.



APPENDIX G. TMPLAR QUICK START GUIDES

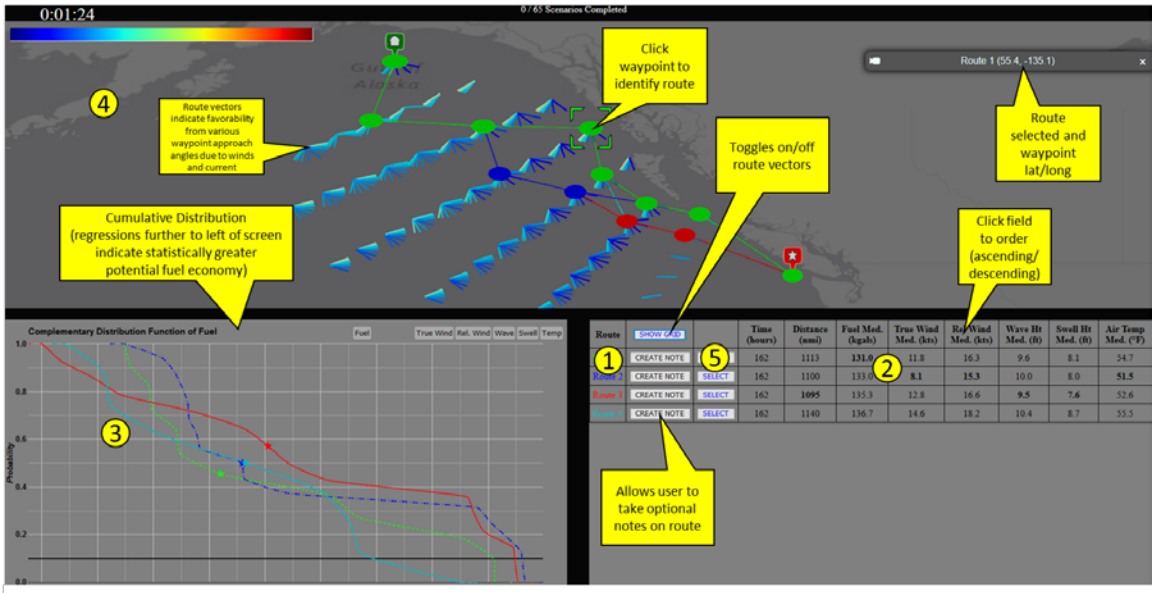
TMPLAR Quick Start (DSS 0)



How to Complete a Scenario:

1. Click on "Route XX" text to display one or more simulated routes on map
2. Compare attributes of calculated routes for maximum fuel efficiency while remaining within safety margins (39' significant wave height, 35 knot true wind speed)
3. Pan and zoom with mouse button/wheel to examine routes on map
4. Click "Select" button associated with the best route to complete scenario

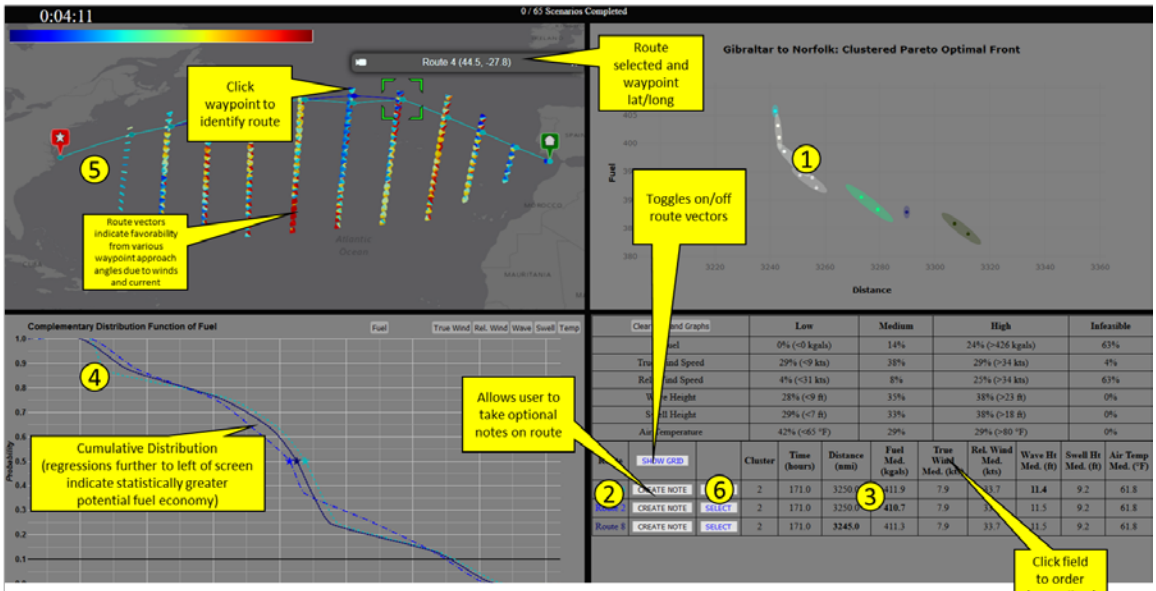
TMPLAR Quick Start (DSS 1)



How to Complete a Scenario:

① Click on "Route XX" text to display one or more simulated routes on map
② Compare attributes of calculated routes for maximum fuel efficiency while remaining within safety margins (39' significant wave height, 35 knot true wind speed)
③ Examine Cumulative Distribution Function for relevant route attributes (regressions toward left of graphic more advantageous)
④ Pan and zoom with mouse button/wheel to examine routes on map
⑤ Click "Select" button associated with the best route to complete scenario

TMPLAR Quick Start (DSS 3)



How to Complete a Scenario:

① Select a group of routes on Pareto Front to populate lower-right quadrant
② Click on "Route XX" text to display one or more simulated routes on map
③ Compare attributes of calculated routes for maximum fuel efficiency while remaining within safety margins (39' significant wave height, 35 knot true wind speed)
④ Examine Cumulative Distribution Function for relevant route attributes (regressions toward left of graphic more advantageous)
⑤ Pan and zoom with mouse button/wheel to examine routes on map
⑥ Click "Select" button associated with the best route to complete scenario

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX H. GAZEPOINT ANALYSIS OUTPUT DATA FIELDS

Gazepoint Analysis Output Data Fields	
Field	**Description**
MEDIA_ID	A unique numeric identifier associated with the media item in the idea list
MEDIA_NAME	The user defined name of a media item
CNT	The COUNT increments once for each data record generated by Control
TIME	Time elapsed in seconds since the last system initialization
TIMETICK	CPU ticks recorded at time as TIME, can be used to synchronize data with other applications
FPOGX	The X-coordinate of the fixation POG, as a percentage of the screen width (0 to 1)
FPOGY	The Y-coordinate of the fixation POG, as a percentage of the screen height (0 to 1)
FPOGS	The starting time of the fixation POG in seconds since the system initialization or calibration
FPOGD	The duration of the fixation POG in seconds
FPOGID	The fixation POG ID number
FPOGV	The FPOG valid flag is 1 for valid and 0 for not valid
BPOGX	The X-coordinate of the un-filtered POG (left & right average), as a percentage of the screen width
BPOGY	The Y-coordinate of the un-filtered POG (left & right average), as a percentage of the screen height
BPOGV	The BPOG valid flag is 1 for valid and 0 for not valid
CX	The X-coordinate of the mouse cursor position, as a percentage of the screen width (0 to 1)
CY	The Y-coordinate of the mouse cursor position, as a percentage of the screen height (0 to 1)
CS	Mouse cursor state, 0 for steady state, 1 for left down, 2 for right down, 3 for left up, 4 for right up
USER	A custom data field that may be set by the user via the API or Remote
LPCX	The X-coordinate of the left eye pupil in the camera image, as a percentage of width (0 to 1)
LPCY	The Y-coordinate of the left eye pupil in the camera image, as a percentage of height (0 to 1)
LPD	The diameter of the left eye pupil in pixels
LPS	The scale factor of the left eye pupil, normalized to 1 at the head depth at calibration
LPV	The left pupil valid flag is 1 for valid and 0 for not valid
RPCX	The X-coordinate of the right eye pupil in the camera image, as a percentage of width (0 to 1)
RPCY	The Y-coordinate of the right eye pupil in the camera image, as a percentage of height (0 to 1)
RPD	The diameter of the right eye pupil in pixels
RPS	The scale factor of the right eye pupil, normalized to 1 at the head depth at calibration
RPV	The right pupil valid flag is 1 for valid and 0 for not valid
BKID	A unique numeric identifier assigned to each blink. Equal to 0 when no blink is detected
BKDUR	The duration of the preceding blink in seconds
BKPMIN	The number of blinks in the previous 60 second period of time
LPMM	The left pupil diameter in millimeters
LPMMV	The left pupil diameter valid flag is 1 for valid and 0 for not valid
RPMM	The right pupil diameter in millimeters
RPMMV	The right pupil diameter valid flag is 1 for valid and 0 for not valid

DIAL	The positoin of the user dial (0 to 1)
DIALV	The dial valid flag is 1 for valid (connected) and 0 for not valid
GSR	The galvanic skin respons value (ohms)
GSRV	The galvanic skin response valid flag is 1 for valid and 0 for not valid
HR	The heart rate value in beats per minute
HRV	The heart rate valid flag is 1 for valid (connected) and 0 for not valid
TTL0	The value of analog input TTL0 channel 0 to 1024
TTL1	The value of digital input TTL1 channel 0 or 1
TTLV	The TTL valid flag is 1 for valid (connected) and 0 for not valid
PIXS	The conversion scale (pixels to mm) if a tracking marker is used
PIXV	The conversion scale valid flag is 1 for valid and 0 for not valid
AOI	List AOI name if the current fixation point is within (overlapping AOIs are hyphenated)
SACCADE_MAG	Saccade magnitude calculated as the distance between the current fixation and last fixation
SACCADE_DIR	Saccade direction calculated as the angle of the vector (current fixation - last fixation) from horizontal

LIST OF REFERENCES

Ahmed, A., & Traore, I. (2007). A new biometric technology based on mouse dynamics. *IEEE Transactions on Dependable and Secure Computing*, *4*(3), 165–179. https://doi.org/10.1109/TDSC.2007.70207

Alter, S. L. (1976). *Computer aided decision making in organizations: A decision support systems typology* [Doctoral dissertation, Massachusetts Institute of Technology]. https://dspace.mit.edu/bitstream/handle/1721.1/47047/computeraideddec00alte.pdf?sequence=1

Boyd, J. R. (1986). *A discourse on winning and losing*. G.T. Hammond (Ed.). Air University Press. https://www.airuniversity.af.mil/Portals/10/AUPress/Books/B_0151_Boyd_Discourse_Winning_Losing.PDF

Boyes, M., & Potter, T. (2015). The application of recognition-primed decision theory to decisions made in an outdoor education context. *Journal of Outdoor and Environmental Education*, *18*(1), 2–15. https://doi.org/10.1007/BF03400975

Brown, J., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. https://doi.org/10.3102/0013189X018001032

Cai, Y., Wen, Y., & Wu, L. (2014). Ship route design for avoiding heavy weather and sea conditions. *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, *8*. http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-04241782-dfb7-49fd-862e-956e6e4dcf0c;jsessionid=F7B4351B97F7591D623EC8752946DCB1

Chang, Y. C., Tseng, R. S., Chen, G. Y., Chu, P.C., & Shen, Y. T. (2013). Ship routing utilizing strong ocean currents. *Journal of Navigation*, *66*(6), 825–835. https://doi.org/10.1017/ S0373463313000441

Chitra, C., & Subbaraj, P. (2010). Multiobjective optimization solution for shortest path routing problem. *International Journal of Electronics and Communication Engineering*, *4*(1), 9. https://waset.org/publications/6469/multiobjective-optimization-solution-for-shortest-path-routing-problem

Chu, P., Miller, S., Hansen, J., & Anderson, B. (2015). Fuel-saving ship route using the Navy's ensemble meteorological and oceanic forecasts. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 12(1), 41–56. https://doi.org/10.1177/1548512913516552

Coyne, J. T., Baldwin, C., Cole, A., & Roberts, D. M. (2009). *Applying real time physiological measures of cognitive load to improve training*. https://doi.org/10.1007/978-3-642-02812-0_55

Ehrgott, M. (2008). Multiobjective optimization. *AI Magazine*, *29*(4), 47–47. https://doi.org/10.1609/aimag.v29i4.2198

Gazepoint. (2019). *Gazepoint analysis user manual.*

Gazepoint. (2020, February 5). *Gazepoint GP3HD product description.* Gazepoint. https://www.gazept.com/product/gp3hd/

Grimes, M., & Valacich, J. (2015). *Mind over mouse: the effect of cognitive load on mouse movement behavior*. https://www.researchgate.net/publication/320471876_Mind_Over_Mouse_The_Effect_of_Cognitive_Load_on_Mouse_Movement_Behavior

Hibbeln, M., Jenkins, J., Schneider, C., Valacich, J., & Weinmann, M. (2014). *Investigating the effect of insurance fraud on mouse usage in human-computer interactions.* https://pdfs.semanticscholar.org/1590/e1a4b3ffa4ad9707295abf30218f173c4a12.pdf

Ide, J., & Schöbel, A. (2016). Robustness for uncertain multi-objective optimization: a survey and analysis of different concepts. *OR Spectrum*, *38*(1), 235–271. https://doi.org/10.1007/s00291-015-0418-7

Ikehara, C., & Crosby, M. (2005). Assessing cognitive load with physiological sensors. *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, 295a–295a. https://doi.org/10.1109/HICSS.2005.103

Joint Chiefs of Staff. (2018a). *Joint maritime operations* (JP 3-32). https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_32.pdf?ver=2019-03-14-144800-240

Joint Chiefs of Staff. (2018b). *Meteorological and oceanographic operations* (JP 3-59). https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/jp3_59.pdf

Joint Chiefs of Staff. (2020). *DOD dictionary of military and associated terms*. https://www.jcs.mil/Portals/36/Documents/Doctrine/pubs/dictionary.pdf

Klein, G. (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Pub.

Klein, G. (2008). Naturalistic decision making. *Human Factors: The Journal of Human Factors and Ergonomic Society*, *50*(3), 456–460. https://doi.org/10.1518/001872008X288385

Klingner, J., Tversky, B., & Hanrahan, P. (2011). Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks. *Psychophysiology*, *48*(3), 323–332. https://doi.org/10.1111/j.1469-8986.2010.01069.x

Krata, P., & Szlapczynska, J. (2018). Ship weather routing optimization with dynamic constraints based on reliable synchronous roll prediction. *Ocean Engineering*, *150*, 124–137. https://doi.org/10.1016/j.oceaneng.2017.12.049

Krejtz, K., Duchowski, A. T., Niedzielska, A., Biele, C., & Krejtz, I. (2018). Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PloS one*, *13*(9). http://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC6138399& blobtype= pdf

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, *14*(5), 331–352. https://doi.org/10.1002/bdecision maker.381

Marcus, G. (2018). *Deep learning: a critical appraisal*. New York, NY: New York University. https://arxiv.org/ftp/arxiv/papers/1801/1801.00631.pdf

Mishra, M., Sidoti, D., Avvari, G., Mannaru, P., Martinez Ayala, D., Pattipati, K., & Kleinman, D. (2017). A context-driven framework for proactive decision support with applications. *IEEE Access*, *5*, 12475–12495. https://doi.org/10.1109/ACCESS.2017.2707091

National Geospatial-Intelligence Agency. (2017). *American practical navigator, volume I* (Pub. No. 9). https://msi.nga.mil/Publications/APN

Olson, D. (2013). Decision support systems. In E. Kessler (Ed.), *Encyclopedia of management theory* (Vol. 1, pp. 185–188). Thousand Oaks: SAGE Publications, Ltd. http://sk.sagepub.com.libproxy.nps.edu/reference/encyclopedia-of-management-theory/i2365.xml

Pan, Y. (2016). Heading toward artificial intelligence 2.0. *Engineering*, *2*(4), 409–413. https://doi.org/10.1016/J.ENG.2016.04.018

Perera, L., & Soares, C. (2017). Weather routing and safe ship handling in the future of shipping. *Ocean Engineering*, *130*, 684–695. https://doi.org/10.1016/j.oceaneng.2016.09.007

Pomerol, J. (1997). Artificial intelligence and human decision making. *European Journal of Operational Research*, *99*(1), 3–25. https://www-sciencedirect-com.libproxy.nps.edu/science/article/pii/S0377221796003785?via%3Dihub

Sidoti, D. (2018). *TMLAR: Tool for multi-objective planning and asset routing* [Unpublished Technical Manual].

Sidoti, D., Avvari, G. V., Mishra, M., Zhang, L., Nadella, B. K., Peak, J. E., Hansen, J. A., Pattipati, K. R. (2017). A multiobjective path-planning algorithm with time windows for asset routing in a dynamic weather-impacted environment. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47*(12), 3256–3271. https://doi.org/10.1109/TSMC.2016.2573271

Shattuck, L., & Miller, N. (2005). *Extending naturalistic decision making to to complex organizations: a dynamic model of situated cognition*. Sage Publishers. https://calhoun.nps.edu/bitstream/handle/10945/47729/Shattuck-Miller-Naturalistic_Decision_Making_2011.pdf?sequence=1&isAllowed=y

Smirnov, A. (2006). *Context-driven decision making in network-centric operations: agent-based intelligent support*. http://handle.dtic.mil/100.2/ADA515001

Sperry Marine. (2014). *Military Integrated Bridge System (IBS)* [Fact sheet]. https://www.electrotech.net.au/wp-content/uploads/2014/01/product-brochure1.pdf

Sperry Marine. (2010). *Voyage Management System VMS VT Professional Integrated Navigation & Bridge Solutions* [Fact sheet]. https://imistorage.blob.core.windows.net/imidocs/8116p075.pdf

Sprague Jr, R. H. (1980). A framework for the development of decision support systems. *MIS quarterly*, 1–26. http://citeseerx.ist.psu.edu/viewdoc/download?doi=1 0.1.1.476.4750&rep=rep1&type=pdf

Susnea, E. (2012). Decision support systems in military actions: necessity, possibilities and constraints. *Journal of Defense Resources Management*, *3*(2), 131–140. https://www.researchgate.net/publication/258833054_Decision_Support_Systems _in_Military_Actions_Necessity_Possibilities_and_Constraints

United States Marine Corps. (2018a). *Command and control* (MCDP-6). https://www.marines.mil/Portals/1/Publications/MCDP%206.pdf?ver=2019-07-18-093633-990

United States Marine Corps. (2018b). *MAGTF meteorological and oceanographic operations* (MCRP 2–10B.6). https://www.marines.mil/Portals/1/Publications/ MCRP%202-10B.6%20Supercedes%20MCWP%203-35.7.pdf?ver=2018-04-05-131541-697

United States Navy. (1995). *Naval command and control* (NDP-6). https://apps.dtic.mil/dtic/tr/fulltext/u2/a304321.pdf

United States Navy. (2011). *United States Navy meteorological and oceanographic support manual* (COMNAVMETOCCOMINST 3140.1M). http://navybmr. com/study%20material/COMNAVMETOCCOMINST%203140.1M.pdf

United States Navy. (2017). *Report on the collision between USS* Fitzgerald *(DDG 62) and Motor Vessel ACX Crystal; Report on the Collision between USS* John S. Mccain *(DDG 56) and Motor Vessel Alnic MC* [Memorandum for Distribution]. Washington, DC: Office of the Chief of Naval Operations. https://s3.amazonaws.com/CHINFO/USS+Fitzgerald+and+USS+John+S+McCain+Collision+Reports.pdf

Vettor, R., & Guedes Soares, C. (2016). Development of a ship weather routing system. *Ocean Engineering*, 123, 1–14. https://doi.org/10.1016/j.oceaneng.2016.06.035

Vogl, J. (2014). *Gazepoint 3 eye tracker manual*. Department of Psychology, University of South Dakota. http://apps.usd.edu/coglab/schieber/eyetracking/Gazepoint/pdf/GP3%20Manual.pdf

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California