NPS-MR-19-002

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

EVALUATING OPERATIONAL EFFECTIVENESS OF DECISION SUPPORT TOOLS

by

Joel Feldmeier, Ta'Lena Fletcher, Mary Jordan, Kurt Klingensmith, Kurt Nielsen, and Brian Wade

February 2019

Approved for public release; distribution is unlimited

Prepared for: U.S. Naval Research Laboratory, Marine Meteorology Division

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

1. REPORT DATE ***(DD-MM-YYYY)*** 22-02-2019	**2. REPORT TYPE** Technical Report	**3. DATES COVERED** ***(From-To)*** December 2018 – January 2019

4. TITLE AND SUBTITLE Evaluating Operational Effectiveness of Decision Support Tools	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 0603207N
6. AUTHOR(S) Joel Feldmeier, TaLena Fletcher, Mary Jordan, Kurt Klingensmith, Kurt Nielsen, and Brian Wade	**5d. PROJECT NUMBER**
	5e. TASK NUMBER 2341
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES) U.S. Army Training and Doctrine Command (TRADOC) Analysis Center (TRAC) Monterey; 700 Dyer Road, Room WA-176, Monterey, CA 93943-0692 Meteorology Department, Naval Postgraduate School; Monterey, CA 93943	**8. PERFORMING ORGANIZATION REPORT NUMBER** NPS-MR-19-002
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Naval Research Lab, Marine Meteorology Division; 7 Grace Hopper Ave, ST2, Monterey CA 93943	**10. SPONSOR/MONITOR'S ACRONYM(S)** NRL MMD
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

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

14. ABSTRACT

This project was a limited look at the effectiveness of meteorology and oceanography (METOC) decision support tools (DSTs). It consisted of a literature search, limited survey of senior officers experienced with both METOC products and being in command of military units, and some suggestions for future work. Several projects deemed to be less burdensome to operational users of METOC products could be started in the near term.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Joel Feldmeier
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UU	51	**19b. TELEPHONE NUMBER *(include area code)*** (831) 656-7776

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ann E. Rondeau President	Steven R. Lerman Provost

The report entitled "*Evaluating Operational Effectiveness of Decision Support Tools*" was prepared for the U.S. Naval Research Laboratory, Marine Meteorology Division and funded by the U.S. Naval Research Laboratory, Marine Meteorology Division.

Further distribution of all or part of this report is authorized.

This report was prepared by:

CDR Joel W. Feldmeier, Ph.D. Department of Meteorology	MAJ Ta'Lena Fletcher TRAC Monterey
Mary S. Jordan Department of Meteorology	MAJ Kurt M. Klingensmith TRAC Monterey
Kurt E. Nielsen Department of Meteorology	LTC Brian M. Wade, Ph.D. TRAC Monterey

Reviewed by:	**Released by:**
Wendell A. Nuss Chair Department of Meteorology	Jeffrey D. Paduan Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This project examined the operational effectiveness of meteorology and oceanography (METOC) decision support tools (DSTs). It consisted of a literature review, a limited survey of senior officers experienced with both METOC products and operational experience, and concludes with suggestions for future work. The literature review focused on best practices and previous work to assess the operational effect of METOC forecasts. The survey results showed that surface wind and seas continue to be important forecasts for military commanders. Synthesizing the best practices for assessment of METOC effectiveness as well as recommended ways forward, this study lays out several suggestions for future work. Several of these suggested projects deemed to be less burdensome to operational users of METOC products could be started in the near term.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION

A. RE-STATEMENT OF PROJECT

This project arose from a desire to collaborate between the U.S. Naval Research Laboratory, Marine Meteorology Division (NRL MMD) and the U.S. Army The Research and Analysis Center Monterey (TRAC Monterey). Personnel from the Meteorology Department at the Naval Postgraduate School (NPS) joined the project to provide additional perspective (including both civilian and military, and operational and academic viewpoints) and subject matter expertise. The authors consider this a pilot project that may lead to more extensive and beneficial studies in the future.

As a part of the Navy's "corporate laboratory", NRL MMD serves the Navy, U.S. Marine Corps, and all of the Department of Defense (DoD) by creating and updating various numerical weather prediction models, Meteorological and Oceanographic (METOC) analyses, and tactical decision aids (TDAs), and/or support or environmental input thereto. As used to help make operational decisions, this report will refer to these products as Decision Support Tools (DSTs). TRAC Monterey provides analytical capability, focused on real-world military problems, with expertise in multiple areas including analysis and visualization of large data, simulation, operations research, leveraging human data sources, and decision-making analysis.

METOC personnel and all others who provide forecasts and/or advice to decision makers need ways to gauge the impacts of their products, to assess when and how to change products, and to decide when to create new products. Although some robust methods exist to verify the accuracy of individual forecasts or forecasts systems, gauging operational impact of those forecasts (or derived products designed to present operational impacts of those forecasts) is much more challenging. In part, this is because the pace of afloat/deployed operations makes detailed recording of decision processes a low priority. Further, METOC impacts are just one of many factors that go into any commander's decision making, and METOC information is discussed and decisions are made at many different times, places, and levels in hierarchies, with or without METOC personnel input, that are never documented. In order to overcome such difficulties, discussions between TRAC Monterey and NRL MMD considered assessing DSTs by analyzing user interactions as possible, as well as a more holistic and operationally focused assessment of how Navy METOC influences missions, i.e., in planning and execution.

B. OVERVIEW OF WORK

The authors and NRL MMD agreed on a three-part approach:

1. Targeted literature review

This consisted of prior studies of operational effectiveness of METOC support by the Navy, U.S. Air Force (USAF), and civilian organizations (both governmental and private, as applicable). The focus was on METOC analyses, forecasts, and impacts predictions. Past results, and past methods for collecting and analyzing data on performance and impacts were of key interest. NPS MR Research Associate Professor Tom Murphree in particular has led several past METOC metrics creation and collection efforts, providing a solid starting point.

2. Survey

The faculty, staff, and students of NPS include active and retired members of all the U.S. military services. As such, they have extensive experience in operations and planning at multiple echelons of command. This includes positive and negative experiences with weather (and ocean and acoustic) conditions on operations, as well as the associated forecasts (or lack thereof) of those conditions. Thus, the survey was designed to assess performance and operational effectiveness of METOC products, as perceived by this (potentially) representative sample of the larger DoD population of decision makers.

3. Roadmap for a future project

NRL MMD and TRAC Monterey decided that initial efforts should take place before the end of calendar year 2018, to help inform NRL MMD discussions with its customers and sponsors in early 2019. Given the short timeline and project start late in 2018, options for future work are included in this report. This report lays out several possible efforts to do in whole or in parts to help NRL MMD with its goals.

II. LITERATURE REVIEW

A. SELECTION OF LITERATURE

Providing relevant information to customers has always been a goal of the larger Weather Enterprise, which includes government, military, and private sector meteorologists. As noted in Katz and Murphy (1997), collaborating with economists and decision science theorists, meteorologists have built a foundation of research into forecast verification, the value of weather information in an economic sense, and ways to apply meteorological information to decision making. The pace of this research accelerated with the announcement of the plan for a Weather Ready Nation (National Weather Service 2011), and since then there has been steady research and demonstration projects involving Impact-Based Decision Support Services (IDSS) in the civilian sector.

Military METOC has always focused on IDSS due to the specific and unique missions of the military. Military METOC also differs in scope and requirements from its civilian counterpart. Therefore, the focus of this initial literature review is on past thesis research at NPS on METOC impacts to operations, and metrics creation and collection. It also included a 2004 NRL technical report, and some civilian aviation weather-decision based research.

B. METOC THRESHOLDS

One way to examine the usefulness of METOC DSTs is by considering METOC thresholds. Weather and ocean thresholds (go/no-go criteria, or red/yellow/green, red/amber/green – corresponding to no-go/marginal/go) have been set for a variety of mission by commanders, at a variety of times. They are supposed to give both METOC professionals and their operational customers pre-determined indications of what forecasts are important and when that forecast is likely to impact operations. One such set of criteria is listed in Appendix A. These are from the 1999 version of the Joint Publication 3-59: Joint Doctrine, Tactics, Techniques, and Procedures for Meteorological and Oceanographic Operations.

The JP 3-59 document and enclosed limits are considered unclassified with no limits on public distribution. While by no means all-inclusive, this is a representative sample of the kinds of thresholds that METOC DSTs as a whole must address. As given, these limits are mission-centric. However, for a METOC science and technology or research and development activity, it is also useful to re-sort the limits by METOC phenomena as is done in Appendix B. For example, rather than listing wind limits under both surface and aviation warfare Appendix B shows all limits within the wind category. This allows researchers to consider if the current forecasts of these phenomena are adequate, and/or if the forecast presentation (e.g., probabilistic instead of deterministic) should be changed or re-negotiated with the customers. Included in Appendix B are some comments on these criteria, how they relate to our literature search, and a note on whether a related metric is seen on the unclassified Fleet Numerical Meteorology and Oceanography Center (FNMOC) metrics page (https://www.fnmoc.navy.mil/verify_cgi/).

C. METOC IMPACTS TO OPERATIONS LITERATURE

1. Safety of operations

The thesis of Martin (2002): METOC and Naval Afloat Operations: Risk Management, Safety & Readiness focused on what might be termed traditional naval weather impacts for ships at sea, or safety in port. He analyzed a set of 8000 reported mishaps to extract 166 of what he termed METOC related mishaps (MRM). Of the MRMs, almost 50% were associated with high winds and seas, and 37% involved rain (primarily in cases of Sailors driving to work). Ice and snow, fog, tides, wind only, and tropical cyclones were also noted, though at much lower frequencies of occurrence. Of particular note, 139 of 166 (84%) mishaps were assessed to involve inadequate training on how to assess, manage, and/or operate in adverse METOC phenomena. In general, there was not enough information to determine what METOC support those involved had received with direct evidence that 8 of 166 and 15 of 166, respectively, did or did not receive a METOC brief.

Cantu (2001) took a similar look at aviation mishaps and weather. Based on record keeping and available data at that time, he focused on 235 Class A (most injurious to personnel or costly in dollars) mishaps that involved aircrew error (from 1990-98). It was assessed that 19% of those mishaps were weather related. By polling experienced pilots on what courses of action they might have taken had they been in the mishap situation, Cantu explored whether if a perfect forecast had been provided to the crew, and the crew believed it, would the accident still have happened? By this methodology, 2/3 of the weather related mishaps were judged preventable. Visibility was a factor in over half of the accidents, including for controlled flight into terrain.

2. Other impacts to operations

Klein (2005) wrote a very interesting technical report, just before the Hurricane Katrina related events in 2005. The report focused on the difficulties in setting Tropical Cyclone Conditions of Readiness (TCCOR; see for example, U.S. Forces Japan 2015) or sorties for Naval commanders in the context of four hurricanes that threatened Florida over a span of several weeks in 2004. Differing storm scenarios related to different decisions-under-uncertainty issues. Klein highlighted the types of uncertainty inherent in the forecast(s), as well as where he assessed that commanders could have benefited from uncertainty information. This report was written before some of the more probabilistic products used by Navy METOC had been promulgated, but the challenges noted echo similar problems that exist today. The author points out that for the sake of forecast continuity, in several cases only small forecast changes were made to pre-set format deterministic products – whereas the qualitative free text forecaster discussions indicated much more uncertainty. Busy commanders typically leave the reading of forecast discussions to their METOC support, and those METOC personnel have limited briefing time. Therefore, useful information about uncertainty was present but likely unused. Klein notes that winds > 50 kts and seas > 12 ft were and continue to be key forecasts for making decision in regards to tropical cyclones. Getting those forecasts correct continues to be a focus of effort for NRL MMD.

Darnell (2006) focused on quantitative assessments of forecasts for the U.S. Air Force and their operational impacts and improving weather support for the war fighter. The particular focus was on Planning Weather Forecasts (PWF), showing that the PWF has a higher potential for making positive contributions to air operations than do Mission Execution Forecasts (MEF). This is important because Air Force Weather (AFW) units spend significantly more time on MEFs than on PWFs. In the case of air operations, the surface visibility, cloud ceilings and cloud layers caused most negative impacts, making these phenomena an important focus for research and training. Results of this study found high levels of mission success even when forecasts were inaccurate, perhaps due to aircrew and mission flexibility. The study revealed a need for improved education of flying units on the nature and availability of AFW products.

3. Metrics creation and collection

Darnell (2006) also tried to expand on immediately preceding work by Butler (2005) for developing METOC impact analyses. Butler developed, tested (and even operationally implemented at NPS), an online tool to ingest METOC data and forecasts provided by the METOC detachment supporting strike and air warfare training and operations for the Navy in Fallon, Nevada. METOC personnel collected and entered METOC data into an open-source data base system at NPS, and a set of statistical tools to evaluate (metrics) the forecasts and data in terms of operations were developed. Hourly weather data collected included observations of Time, Wind speed, Visibility, Weather condition, Sky condition (cloud density and heights), and combined seas, as well as forecasts provided. Observations were converted into common categories of Red, Yellow or Green in terms of metrics developed by other NPS Meteorology thesis students. The goal was to quantify Forecast Accuracy (FAC), Probability of Detection (POD) and Number of Accurate Forecasts (NAF). An intended product of this effort was a final metrics report format that supports the needs of users (i.e., the METOC units and their operational customers), as users have determined as their needs. This does highlight the need for extensive customer involvement

Callahan (2006) continued the work of Butler and Darnell. He proposed several metrics to properly evaluate forecaster’s capability, grouped into three broad categories: forecast performance metrics, operational impacts metrics, and impacting phenomena metrics. Forecast performance metrics (forecast accuracy metric, probability of detection metric, false alarm rate metric and bias metric) assess the quality of the forecast. Operational impacts metrics included received negative metric, mitigated received negative metric, missions placed at risk metric, mitigation rate metric, missions requiring mitigation metric, missions canceled metric and targets change metric. They describe “how information in the forecast and the actual weather phenomena experienced during the mission affected or could have affected operations.” Impacting phenomena metrics include mission impacting phenomena metrics, mission canceling phenomena metrics and target changing phenomena metrics. “These relate individual weather phenomenon to forecast performance and impacted operations.”

Callahan also highlighted some important considerations:

1. The timely arrival of forecast materials. Sometimes a forecast may be too late for mission planning purposes but not mission execution.
2. Having weather personnel present at mission briefings greatly benefited operators because they were able to get a detailed explanation of the atmospheric conditions and could ask questions as to how those conditions could affect their mission. This provided the operators with the necessary information to pre-plan contingency plans for varied weather.

The metrics developed by Callahan were used to create a system to provide near real time metrics reports. The program went through several revisions until it reached a full-scale test implementation online. The work of Butler, Darnell, Callahan, and others under Professor Murphree at NPS determined that a metrics system is viable. Implementing this system does require a significant “buy-in” from both deployed METOC personnel and their supported customers. This is the biggest barrier to its successful implementation. Darnell in particular found difficulty in getting pilots to provide information post-mission. If the NPS metrics system were to be re-invigorated and expanded, Callahan had several suggestions that still apply today:

1. Development of a comprehensive METOC metrics program directed by CNMOC.
2. The adaptions made to prior studies such as Darnell (2006) could be used in the sister services to refine their METOC capabilities.
3. Development of a metrics system that combines mission planning and execution forecasted data.
4. Continual integration of METOC data collection efforts. Being able to extract the data from multiple systems in an automated process without the need of outside assistance to be used in a METOC metrics program is beneficial to the community as a whole.

4. METOC and decision making

The Federal Aviation Administration (FAA) sponsored Robinson, et al., (2011) to examine the implementation of high-resolution deterministic forecasts in concert with procedural/coordination changes. Of specific concern was convective weather forecasts and air traffic control. The new system was intended to assist what were termed strategic air traffic decisions. For the FAA's purposes, strategic referred geographically to the whole national airspace, and temporally to weather constraints 2-8 hours ahead of time. In support of this time scale, an 8 hour "radar-like" forecast was made available at FAA and airline dispatch facilities, with updates every 15 minutes at an ~3 km scale. Two items particularly stood out from the FAA study, in regards to NRL MMD's desire to assess METOC product usefulness:

1. The product had been tailored to the customer needs i.e., using radar return products to understand thunderstorms is something air traffic controllers have seen before. Therefore, the new high-resolution forecast information was in that format. The lesson being: present a customer with forecast information in as close a format as to what they will experience operationally. The challenge can be to find that format.

2. In order to do a rigorous assessment, Robinson, et al., had to coordinate multiple "observation blitzes". These included observers, well trained in the new forecast products and with familiarity with air traffic control, to be stationed "looking over the shoulder" of air traffic controllers in real time, assessing how the new product was used. This was a significant investment of observer effort as well as a potential distraction to the customer.

An earlier, but revealing project is that of Lind, et al., (1994). This study illustrates a process for targeted presentation of METOC information. In this case, MIT Lincoln Laboratory, through the sponsorship of the FAA, was developing a data link application to provide graphical weather information to the general aviation pilot in the cockpit. As part of the design process, a human factors study was undertaken to understand the impact of Graphical Weather Service (GWS) images on pilot actions.

Twenty instrument-rated pilots participated in five hypothetical flights, one for training and four for data collection. Ten of the pilots had moderate instrument flight rules (IFR) experience and ten if the pilots had extensive IFR experience. Each pilot had four hypothetical flights: two with GWS images and two without GWS.GWS images were available to the subject during the training flight and two of the four data collection flights. Subjects were provided with relevant navigational charts and weather briefing material prior to each "flight." At three decision points within each flight, the experimenter told the subject the aircraft position and altitude, described the current weather conditions in the immediate vicinity of the aircraft, and asked the subject what action he or she would take. The subject could respond immediately, or seek additional information using the Graphical Weather Service (GWS) or via queries to weather dissemination personnel (an experimenter, who sat in the room with the subject, played that role).

Pilots were asked to "think aloud" throughout the hypothetical flights. The experimenter recorded the subject's choices of GWS images, queries to weather dissemination personnel, and all comments made at each decision point. In addition, the subject completed a questionnaire at each decision to assess his confidence in his ability to assess the weather situation and to assess the usefulness of GWS as an aid in the decision-making process. Weather for each flight was different, but each included convection and/or threat of thunderstorms. Locations of strong convection, precipitation, visibility and ceilings changed during the flight. Pilots chose flight path deviations as needed. From written or verbal in-flight weather updates, the precise location of convection and precipitation was not known. The forecast of “thunderstorms may form in the vicinity” did not provide specific convection location or timing. GWS radar images provided precipitation location information and GWS used three colors to depict the echo intensity levels for weak, moderate and strong-to-extreme precipitation. The pilot was to infer the potential for turbulence from the radar images.

Lind, et al., assessed that GWS had a substantial positive effect on the weather-related decisions made by the subjects. With GWS, subjects could see the weather graphically displayed and could make informed decisions regarding whether to embark on a flight and regarding the need for deviations for weather avoidance during flight. Subjects' confidence in their ability to assess the weather situation was significantly increased when GWS was used. Subjects found GWS to be very useful and cost-effective. Pilots were enthusiastic about receiving graphical weather information in the general aviation cockpit. There was no statistical difference between pilot responses depending upon pilot experience level. Data obtained from subject comments and the observations of the experimenters indicate that GWS was not used as the sole source of information in making a decision. GWS was used to confirm, clarify, and augment information the pilots had from other sources.

In a 2006 thesis, Cunningham used air-refueling scenarios to demonstrate the integration of probabilistic turbulence forecast guidance into the U.S. Air Force operational risk management (ORM) process. The overarching idea from prior studies was that if a user's exposure to risk can be quantified and related to their risk tolerance, then better (more economical) decisions can be made. The associated Cost-Loss (C-L) ratio for the user of a forecast can be difficult to determine in DoD operations. Therefore, instead of C-L, some forecast users make decision based on a more qualitative assessment through operational risk management (ORM).

Mission planners may be able to describe C and L in terms of benefits and risks. For example, military planners balance risk with mission priority. Cunningham argues that for air-refueling scenarios with high and low C-L user ratios, ensemble-based probabilistic forecasts provided more value at Tau-24 and Tau-48 valid times than deterministic forecasts valid at corresponding times. For scenarios with middle-range C-L ratios, probabilistic and deterministic forecasts provided nearly-equal value at Tau-24, but probabilistic forecasts provided more value than deterministic at Tau 48.

Szczes (2008) continued ensemble related and ORM based assessments. Her thesis attempted to account for uncertainty (ambiguity) in both the ensemble data and user risk tolerance for the decision input. Ambiguity was defined as the uncertainty of the uncertainty prediction, which can be conceptualized as errors bars about a probability forecast. Alternatively, ambiguity can be considered as the potential error in the ensemble's forecast probability distribution function. To incorporate uncertainty into both the ensemble forecast and risk tolerance, the researcher created a decision input visualization tool for varying levels of risk intolerance, mission thresholds for turbulence, flight levels, and forecast hours. When applied to forecasting clear air turbulence for aircraft refueling flights, a deterministic output of the decision aid identifies large regions of "inconclusive" – where the results is neither acceptable risk (Go) or unacceptable risk

(No Go). Using an ensemble-based decision aid, which accounts for ambiguity, the regions of inconclusive decisions were dramatically reduced.

Palmer (2010) also looked at ensemble based probability. His thesis explored the value of stochastic forecasting through a series of operational events, in the context of a Strike Warfare campaign in the Weather Impact Assessment Tool (WIAT), a campaign simulator. Unfortunately, limitations of the WIAT tool prevented conclusive results and post-2010 information about the WIAT tool could not be found in literature or web searches.

THIS PAGE INTENTIONALLY LEFT BLANK

III. SURVEY METHODOLOGY AND RESULTS

Appendix C lists the survey questions used in this study. NPS uses the LimeSurvey tool for these surveys (https://www.limesurvey.org/).

A. DESIGN OF QUESTIONS AND SELECTION OF PARTICIPANTS

As mentioned in the introduction, the population of NPS includes individuals presumed to have familiarity with METOC products and personal as well as command or operations/planning experience based on their current seniority and positions. This population has made operational judgements throughout their careers, both with and without adequate METOC input, as junior officers and in positions that are more senior. Senior officers (O-5 and above), on active duty or retired, from the Army, Navy, Air Force, and Marines from the faculty and staff were sought out for this initial survey.

The opening questions of the survey regarded familiarity with the thresholds listed in an older version of the Joint Publication 3-59 (Joint Chiefs of Staff 1999). Note that all participants received a copy of these thresholds to read as shown in Appendix A. Although this particular set of thresholds is incomplete, this set covers a representative sample of air, sea, and land operations. The intent of such thresholds is for operational customers to tell METOC personnel what is important for a particular operation. Therefore, survey participants were asked if they agreed with these representative parameters, and if they received forecasts or products that included those parameters.

The survey also included questions on how recently the participants had dealt with METOC products, and at what level of command they were at that time. Considering the tactical, operational, and strategic levels of warfare, the space and time scales a commander typically cares about grow with seniority. For example, a senior aviator would have cared most about short term forecasts of wind, icing, cloud cover, etc., as a junior pilot planning a mission, just hours in advance. Years later, that same but more experienced pilot may be involved in planning an operation a week in advance and would desire products in a different format. Such a range of experience informs the answers for the whole survey.

As noted in the Commander, Naval Meteorology and Oceanography Command Strategic Plan (2018), operational relevance is critical, as is figuring out the best ways to make use of probabilistic forecasts and pushing forecasts further into the future. Therefore, participants were asked about perceived usefulness, timeliness, and accuracy, as well as agreement or disagreement that METOC is important to planning. The survey also asked for input from the participants on if uncertainty information is useful, would they take a longer lead forecast even with greater uncertainty, and how much uncertainty they could tolerate for their operations. To close the survey, there were questions on the most and least useful METOC products, how much METOC training a participant might have received, and a chance to provide anecdotes or complaints.

B. RESPONSE RATE

Some of the authors informally socialized the idea for such a survey to a few senior officers on campus prior to the project formally starting, and received positive feedback. However, as of the time of this writing only 8 people had responded, with only 5 complete responses. This was likely due to the timing of when the survey was distributed during the academic year. The final Institutional Review Board (IRB) approval was complete and the study team was able to send the survey in late December. Classes and exams were completing, and most at the school were preparing for some form of holiday leave. The team has re-contacted the invited survey participants, but this report is being written before new results have come in. The team will deliver an updated survey analysis to NRL MMD if it is able to get more responses, and if the new responses are particularly insightful.

Of the five complete responses, three were from Navy Captains (O-6), one was an Army Colonel (O-6), and one had been a Major (O-4) in the Marine Corps. All but one had more than 20 years of active duty experience, and all but one had command experience (one at the O-5 level, and three at the O-6 level). In regards to training, all professed to have had formal meteorology training during their careers, two to having oceanography training, and none to having space weather training.

C. SUMMARY OF RESPONSES

As mentioned, the Joint Publication 3-59 is only one possible source of METOC thresholds. However, it was interesting that three of our five respondents were not familiar with that publication. As far as the forecasts that they have previously received being consistent with the thresholds, four of the five indicated that they were. One respondent disagreed with the given limits for ship underway replenishment, saying that 16 ft seas would be “red” in his or her opinion. Another respondent suggested multiple additional parameters for anti-surface warfare and anti-submarine warfare, although these may be present on other threshold lists. For flight operations, one respondent commented that ceiling limits are very mission dependent.

In regards to lead times for products, the participants in this sample indicated that between 6 – 24 hours was sufficient. However, they would be open to increasing lead times for mission planning purposes. The scope of this lead-time increase is, of course, mission dependent. From a planning perspective, the respondents all indicated that if the forecast was not accurate at 72 hours, mission degradation is experienced.

As to what they perceived as the most useful products, there was little commonality among the participants. The products mentioned were:

1. Optimal Track Ship Routing (OTSR) and Operating Area (OPAREA) Forecasts
2. Standard Flight Weather Briefings (DD-175-1)
3. Real time updates on changing or deteriorating weather conditions
4. Weather (winds, visibility and ceilings)
5. Passive acoustic predictions for threat submarine frequencies

In regards to least useful products, those listed were:

1. General Forecasts
2. Weather updates associated with areas outside of the area of responsibility (AOR) that didn’t include a divert airfield
3. Times for nautical dusk and dawn (EENT/BMNT)
4. Weather report in port

The ending question of the survey asked for final thoughts. Again there was little commonality, consistent with such a small sample. One suggested changes to the provided thresholds to include more laser related parameters. In-person METOC assistance was noted to provide the best consumer experience, allowing for explanation of products. Better real-time communication of weather updates to fixed wing aircraft, surface winds around target area to assist Direct Air Support, working on radar capability to pick up clouds at mid to low levels, and creating three-dimensional views of electromagnetic/electro-optical sensor effectiveness were also mentioned.

IV. POSSIBLE FOLLOW-ON EFFORTS

A robust examination of METOC product effectiveness on commander decision-making, and associated outcomes, requires extensive data. Ideally, a future researcher would have unfettered access to observe decision-making at multiple echelons of command. In particular, the study team would especially want to see what sorts of METOC information a commander receives, the format of that information, and the other factors that a commander must take into account. However, it is rare that an operational command has time, space, patience, and ability to deal with extra and external personnel solely focused on metrics, engaged in intensive metrics collection. Therefore, the recommendations below are divided into those that might have low negative impact on a customer or commander, and those that would be more intrusive. As previously noted, a series of NPS thesis students under the guidance of Professor Murphree developed a prototype metrics system that would be an excellent starting option. This would require significant customer "buy-in", at least to change standard operating procedures to include more data collection.

A. LESS INSTRUSIVE EFFORTS

1. Operational chat records

Starting in the late 1990s and continuing until the present, the use of online chat programs for coordination and command and control has continuously grown across the military (Heacox 2004). Commands may keep records of chat exchanges, and these records present an intriguing source of useful information. Although an individual commander might not be in a chat room his or herself, their direct representatives typically are. The various chat sessions are a place where lower echelons can report METOC related problems or conditions to their bosses, and begin discussions on possible courses of action. Additionally, METOC commands and deployed METOC personnel use chat for coordination, de-confliction of forecasts, requests to subject matter experts, mutual advice, and general discussion.

Chat records are ripe for text analytics. With carefully selected keywords, TRAC Monterey and collaborators may be able to extract and sort from records trends in the mention of various METOC phenomena, and/or their operational impacts. Follow-on analysis would aim to identify the types of METOC related questions operational personnel ask. This would allow for such things as: 1) assessment of current METOC products and if they offer the kind of information desired; 2) if the information desired is available, do operational personnel understand the products they get or the briefings they receive; and 3) are there gaps between customers desires and METOC's ability to deliver a product. Records might also expose persistent misunderstandings by supported warfighters or METOC personnel. Chat records are also valuable because they represent thinking and (virtual) conversations "in the moment", and capture detail and nuances that might be forgotten in later debriefs. The records could capture situations that deployed METOC personnel or supported warfighters meant to report to someone, but did not due to distractions from operational pressures.

As a first step, TRAC Monterey could data-mine all available historical files. This would help develop appropriate tools and keyword, and reveal past trends. Then as a follow-on, working with NRL Monterey, a real-time or near real-time tool might be possible. Depending on chat availability, an application could be made that would alert METOC watch-standers of questions and/or problems with METOC analyses, models, forecasts, or other products. For example, such an application might flag mentions of "limits" or "weather holds". Ongoing analytics, possibly using artificial intelligence based techniques, could help reveal trends in negative METOC impacts to operations and perhaps even link with more traditional forecast/model metrics (see for example https://www.fnmoc.navy.mil/verify_cgi/).

2. Logic models

A recent study in support of the National Weather Service's Weather Ready Nation (WRN) initiative suggested the use of logic models (Nadeau 2017). In this case, logic models are simplified flowchart depictions of how a forecast provider might think a customer would use their forecast. As a very simple example, an afloat Navy forecaster may think the decision for flight operations may look something like:

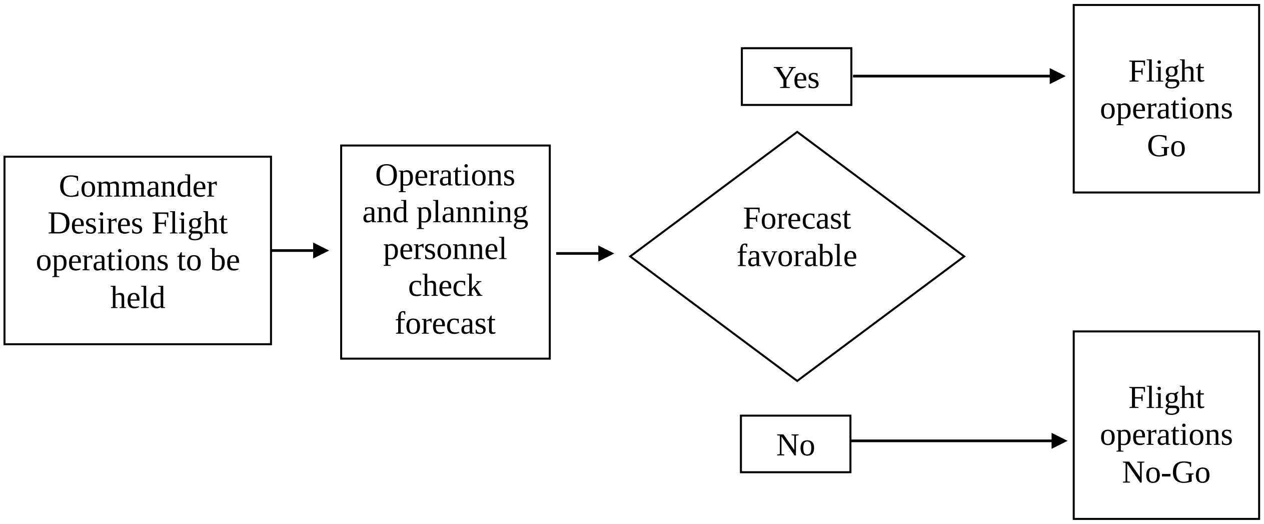

Figure 1. Simplified Logic Model for Flight Operations Decisions

However, multiple other factors may come into play. For example, lower echelons might ask for maintenance time instead of flying, or there might be operational reasons to fly or not fly unrelated to weather. In addition, an unfavorable forecast might lead to inserting steps into the decision process to check, several hours before the flight operations, if a forecast changes or if observations are matching the forecast – before finally cancelling flight operations.

TRAC Monterey could collaborate with NRL MMD, along with personnel from Commander Naval Meteorology and Oceanography Command to create a base set of logic models for how that group assumes forecasts or METOC products are used. Then, the study would ask METOC personnel recently returned from deployment to modify these naïve models to “how it was really done” on their deployment. Personalities and preferences of commanders and METOC personnel vary from unit to unit and deployment to deployment, so likely the end result would be an ensemble of logic models of how METOC information is used. Researchers would crosscheck these ensembles with the suggested chat log analyses. This could allow for updated and modified logic models. This might help NRL MMD and others to modify content of products, delivery timing, and distribution of METOC information.

Large deviations from the naïve logic models may reveal when customer training may pay dividends in introducing new products, or better explaining and/or enhancing the use of currently ignored products. Validated logic models across multiple mission areas could also be a valuable tool to show the return on investment (ROI) from METOC efforts.

3. Additional analytics

In addition to chat logs, multiple lessons learned systems exist within the Navy and the DoD. Although a valuable resource, these are not all in the same format. Though searchable by keyword, in general they are not easy for humans to sort through unaided due to nuggets of information, for example, buried within multiple types of reports on multiple events. In addition, as noted in the literature review, past examinations of Navy Safety Center aviation accident records revealed METOC related trends. Since that study, tracking of METOC related impacts to operations has improved. A recent inquiry to the Navy Safety Center yielded information on 846 accidents with some relation to METOC, from 2010 through late 2017 (K. Perry, personal communication, August 15 2017). The study team could apply text analytic techniques to all of these sources of information.

Additionally, as previously proposed by TRAC, it may be useful to do an assessment of all DSTs supported by NRL MMD, be they hosted at NRL MMD or at the Fleet Numerical Meteorology and Oceanography Center (FNMOC). TRAC and partners could assess the workflow and use of the DSTs as a whole. The study team would also focus on one or two specific DSTs for more in-depth analysis. The goal would be to gain insights into which tools are most used; how, when, where, and by whom are the tools used; and the level of interaction required by the user to obtain their final products.

B. MORE INTRUSIVE EFFORTS

As previously described, Robinson et al. (2011) used an “observation blitz” approach to track the usage, interpretation, and performance of a newly introduced product (a weather planning tool, by the FAA, for air-route planners). In that case, highly trained observers were “present at several FAA and airline facilities…” and “… observations at each facility were made simultaneously in order to better understand the coordination and collaboration interactions associated with strategic weather impact mitigation planning.” Such intense efforts have also occurred within the last 15 years within Navy METOC (J. Feldmeier, personal communication, 2019). However, they are cumbersome to set up, e.g., arranging travel to and berthing on multiple, already crowded, vessels for highly skilled personnel, who are not present to aid the observed unit's mission. Such Naval efforts have also focused on the introduction of experimental products, which given operational pressures an intended customer may push aside in favor of existing products they are more comfortable with. Therefore, any naval METOC observation blitz should be rare and very specific. However, other data gathering options exist.

1. Surveys

Although the study team received limited responses from the survey at NPS, this was likely due to the timing (i.e., the survey was approved and released days before Christmas, leading to invited participants overlooking or forgetting the survey). The general design of the survey, both kinds of questions and intended audience (commanding officers, former commanding officers, and/or those with significant operations and planning experience) should provide useful data. A second effort could explore the procedure and permissions required to invite a much larger audience to the survey. For example, the study team could invite large numbers of personnel stationed at the Pentagon to participate. Alternatively, the team could invite large segments from the shore commands (presumably staffed by personnel with recent relevant operational experience, who since not deployed would have time available to give a survey serious thought) associated with Naval Aviation, Special Warfare, the Submarine Force, the Surface Force, Expeditionary Forces, and others. Future work should also examine U.S. Marine Corps needs, as well as seeking out joint examples. Surveying the (more junior) student population at NPS is also an option.

2. Active monitoring and polling

If not already active, develop across all domains an automated complaints / comments / suggestions data gathering application. For example, many METOC products have text to identify where they were created, and include a forecaster's name with some sort of contact information. When a customer clicks on such a "problem" or "contact" link, rather than simply opening Microsoft Outlook for an email to a watch stander, either an additional metrics collection email could be auto-included or the customer could be directed to a comment system like those typically used by information technology support centers. Properly constructed, such a system would allow for text analytics, among other techniques, to help sort data, look for trends in problems, complaints or questions, and indicate areas for improvement.

Taking inspiration from the business world, if information assurance policies allow it, one could also randomly ask customers for voluntary feedback. For example, for every twentieth unique user to access tsunami-warning graphics from the Joint Typhoon Warning Center, a pop-up window or some other interaction could occur asking that customer to leave a comment or answer a brief survey. This would have to be voluntary, and designed in such a way as to have minimal impact and annoyance for the customer.

3. Improved tools and procedures for deployed METOC Personnel

NRL MMD supports the Navy in general, but often through deployed METOC personnel. Thus, the customer, or “the warfighter” for deployed METOC is usually pilots, SEALs, ship captains, admirals, etc. For NRL MMD, sometimes the terms “customer” or “warfighter” evokes the deployed (or on operational watches ashore) METOC personnel (J. Hansen, personnel communication, 2017). There is an opportunity, however, for NRL MMD and deployed METOC, with direction and coordination with CNMOC, to develop and improve data collection and metrics to measure METOC effectiveness. As a simple example of a process improvement, consider the following based on the heavy use of email underway for a Strike Group Oceanography Team (SGOT), deployed on a carrier or large-deck amphibious ship:

First, train watch standers and supervisors to save an (ideally) shared email archive file. This would contain all service requests, questions, and complaints they received by email at that watch station. Also, train the watch to forward to that archive any existing pass-down log documents, or a summary email written by them (once a watch) of customer interactions they had in person or by phone. METOC personnel who provide direct staff or planning support, such as a strike forecaster or a strike group oceanographer, could also be trained send to that shared email a summary of their daily activities. Then daily, weekly, monthly, and at the end of a deployment or exercise run text analytics on that email archive, again to help expose trends. Although deployed METOC might find such tasking annoying at first, and likely the record would be incomplete, with time such a procedure would become a normal part of underway work and presumably provide useful insights.

Additional methods of automated data collection are possible. However, in the near term, NRL MMD may be able to collaborate with an existing program. Over the past few years, although FNMOC is not a deploying command, the FNMOC Reach Back Cell occasionally sends subject matter experts to meet units on deployment to help them better understand available METOC products and procedures. On a not to interfere basis with their FNMOC tasking, these temporary ship riders could observe or collect metrics on how METOC products are used. Admittedly, it would likely take several attempts to come up with a reasonable (and small) set of questions, requests, or observations for these personnel to make. FNMOC is willing to discuss possibilities (T. Keefer, personal communication, 2019).

THIS PAGE INTENTIONALLY LEFT BLANK

V. SUMMARY

It is not always straightforward to measure METOC effectiveness. From the standpoint of what METOC parameters to focus on, Appendix B provides a good starting point. From surveys of senior military officers, surface wind and seas continue to be important. Training customers to write their thresholds probabilistically would be of interest. Various flight parameters such as ceiling, cloud cover, icing, and visibility also remain challenging, as do thunderstorms. Particular attention should be paid to presenting products in formats and with a level of precision that meets customer needs – or if the precision requested is not realistic, more discussion with customers on what is possible should be held.

To understand if a forecast is useful, the study team would have to be able to assess if it is correct and how it is used. The academic and modeling communities have understandably focused on metrics for gauging quantitative correctness of numerical forecasts. However, the various ideas examined above for such things as operational impacts metrics and impacting phenomena metrics have merit and should be re-examined. As Callahan (2006) pointed out, some of this should be done at the CNMOC level with NRL input. In fiscal year 2019, NRL MMD should continue to collaborate with TRAC on implementing some or all of the suggested “less intrusive” options, and consider future collaboration for larger efforts.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A

Reproduction of Appendix G, Joint Publication 3-59 (1999)

METOC IMPACTS TO SELECTED OPERATIONS

OPERATION	FAVORABLE (No Degradation) (GREEN)	MARGINAL (Some Degradation) (AMBER)	UNFAVORABLE (Significant Degradation) (RED)
ARFOR OPERATIONS			
BRIDGING	WIND < 10 KTS	WIND 10-34 KTS	WIND > 34 KTS
ARMOR GUN SIGHTS	VIS > 2000 m	VIS 1000-2000 m	VIS < 1000 m
TOW MISSILE	VIS > 3000 m	VIS 2000-3000 m	VIS < 2000 m
HELICOPTER (LIFT)	CIG > 500 FT	CIG 300-500 FT	CIG < 300 FT
(no specific airframe)	VIS > 1600 m	VIS 800-1600 m	VIS < 800 m
	NO ICG/TURBC	LGT OR MDT	SVR TURBC/ICG
		TURBC OR ICG	
HELICOPTER (ATTACK)	CIG > 2600 FT	CIG 1100-2600 FT	CIG < 1100 FT
(no specific airframe)	VIS > 4000 m	VIS 1000-4000 m	VIS < 1000 m
	WIND < 25 KTS	WIND 25-50 KTS	WIND > 40 KTS
			TEMP > 90F
	NO PRECIP	MDT PRECIP	HVY PRECIP
	NO Thunderstorms	FEW Thunderstorms	SCT Thunderstorms
HELLFIRE MISSILE	CIG > 2000 FT	CIG 800-2000 FT	CIG < 800 FT
	VIS > 5000 m	VIS 3000-5000 m	VIS < 3000 m
CLOSE AIR SUPPORT	CIG > 2000 FT	CIG 1000-2000 FT	CIG < 1000 FT
(For Army Planning Purposes)	VIS > 8000 m	VIS 3200-8000 m	VIS < 3200 m
AERIAL RECON	< 2/8 CLD COVER	2/8-4/8 CLD COVER	> 4/8 CLD COVER
	VIS > 8000 m	VIS 4800-8000 m	VIS < 4800 m
	Aerial Recon covers three levels — Strategic (above 25,000 ft), High (8,000 - 25,000 ft), and Low (below 3,000 ft). Cloud cover is for at or below flight (operating) level.		
GROUND RECCE	VIS > 3000 m	VIS 1000-3000 m	VIS < 1000 m
PARADROP	WIND < 13 KTS	WIND 13-18 KTS	WIND > 18 KTS
			CIG < 1000 FT
		LGT PRECIP	MDT/HVY PRECIP
	DA < 4000 FT	DA 4000-6900 FT	DA > 6900 FT
	(DA = Density Altitude)		
NBC OPERATIONS		WIND < 10 KTS	WIND > 30 KTS
			WIND CALM
			Wind From Enemy
	Stable Atmosphere	Neutral Stability	Unstable Atmosphere
	NO PRECIP	LIGHT PRECIP	MDT/HVY PRECIP
SMOKE		WIND 5-10 KTS	WIND < 5 KTS
			WIND > 19 KTS
			Wind From Enemy
			TEMP > 120 F
		MDT PRECIP	HVY PRECIP

METOC IMPACTS TO SELECTED OPERATIONS (cont'd)

OPERATION	FAVORABLE (No Degradation) (GREEN)	MARGINAL (Some Degradation) (AMBER)	UNFAVORABLE (Significant Degradation) (RED)
ARFOR OPERATIONS (cont'd)			
PERSONNEL	NO PRECIP	LIGHT PRECIP	MDT PRECIP
(temp — heat and/or wind chill indices)	> 20F and < 85F	85-95F or -15 to 20F	> 95F or < -15F
Lock On Before Launch	CIG > 1900 FT	CIG 400-1900 FT	CIG < 400 FT
	VIS > 7000 m	VIS 500-7000 m	VIS < 500 m
Lock On After Launch	CIG > 1700 FT	CIG 800-1700 FT	CIG < 800 FT
	VIS > 7000 m	VIS 1700-7000 m	VIS < 1700 m
COPPERHEAD	CIG > 3000 FT	CIG 1000-3000 FT	CIG < 1000 FT
	VIS > 2500 m	VIS 1000-2500 m	VIS < 1000 m
	NO PRECIP	MDT PRECIP	HEAVY PRECIP
SEA PORTS	WIND < 20 KTS	WIND 20-35 KTS	WIND > 35 KTS
AIR PORTS	CIG > 1500 FT	CIG 200-1500 FT	CIG < 200 FT
	VIS > 4800 m	VIS 900-4800 m	VIS < 900 m
AIR DEFENSE	CIG > 5000 FT	CIG 2500-5000 FT	CIG < 2500 FT
	VIS > 5000 m	VIS < 5000 m	
			TEMP > 120F
ARTILLERY FIRES	CIG > 1500 FT	CIG 600-1500 FT	CIG < 600 FT
	VIS > 3000 m	VIS 1000-3000 m	VIS < 1000 m
	WIND < 30 KTS	WIND 30-35 KTS	WIND > 35 KTS
			TEMP < 20F
			TEMP > 125F
	NO TO LGT PRECIP	MDT PRECIP	HVY PRECIP
SIGINT	WIND < 30 KTS	WIND 30-45 KTS	WIND > 45 KTS
		TEMP 85-120F	TEMP < 32 F
			TEMP > 120F
TRAFFICABILITY	NO PRECIP	MDT PRECIP	HVY PRECIP
AFFOR OPERATIONS			
AIRLIFT	CIG > 1000 FT	CIG 500-1000 FT	CIG < 500 FT
(no specific airframe)	VIS > 8000 m	VIS 4800-8000 m	VIS < 4800 m
		Light Freezing Precip	Freezing Precip which closes runway
FLIGHT OPERATIONS	CIG > 3500 FT	CIG 1000-3500 FT	CIG < 1000 FT
(CAS and/or DEEP ATTACK)	VIS > 3200 m	VIS 1600-3200 m	VIS < 1600 m
(no specific airframe)	WIND < 25 KTS	WIND 25-35 KTS	WIND > 35 KTS
		LTG-MDT TURBC	SVR TURBC
ELECTRO-OPTIC SUPPORT	CIG CLR-SCT	CIG BKN	CIG OVC
(Absolute Humidity Limitations)	< 14 g/m3	14-18 g/m3	> 18 g/m3
(Transmittance)	> .4	0.2-0.4	< 0.2
(Moon Illumination)		Moonrise/Moonset	No Moon
(no specific system)			HVY PRECIP

METOC IMPACTS TO SELECTED OPERATIONS (cont'd)

OPERATION	**FAVORABLE (No Degradation) (GREEN)**	**MARGINAL (Some Degradation) (AMBER)**	**UNFAVORABLE (Significant Degradation) (RED)**
AFFOR OPERATIONS (cont'd)			
PREDATOR	CIG > 2000 FT	CIG 800-2000 FT	CIG < 800 FT
	VIS > 4800 m	VIS 3200-4800 m	VIS < 3200 m
	Crosswind < 10 KTS	Crosswind 10-15 KTS	Crosswind > 15 KTS
AIR REFUELING	NO CLOUDS AT FLIGHT LEVEL	SCT-BKN CLOUDS AT FLIGHT LEVEL	OVERCAST CLOUDS AT FLIGHT LEVEL
	NO Thunderstorms	FEW Thunderstorms	SCT Thunderstorms
NAVFOR OPERATIONS			
AMPHIBIOUS WARFARE	CIG > 5000 FT	CIG 300-5000 FT	CIG < 300 FT
	VIS > 4800 m	VIS 1000-4800 m	VIS < 1000 m
LANDING CRAFT			
(Combined Seas)			> 5 FT
(Percent Illumination)			> .0001 FT CANDLES
(Moderate Surf Index)	< 8	8-10	> 10
(Breaker Heights)	< 5 FT	5-12 FT	> 12 FT
(Wake Period)	> 7 SECONDS	6-7 SECONDS	< 6 SECONDS
ANTI-SURFACE WARFARE			
Over the Horizon — Targeting	NO/LGT PRECIP	MDT PRECIP	HVY PRECIP
			TEMP > 103F
(Seas)	< 6 FT	6-8 FT	> 8 FT
			WIND > 60 KTS
COMBAT SERVICE SUPPORT			VIS < 400 m
(Combined Seas)	< 12 FT	12-19 FT	> 19 FT
MINE WARFARE AVIATION			CIG < 300 FT
			VIS < 1000 m
	WIND < 25 KTS	WIND 25-35 KTS	WIND > 35 KTS
EOD DIVERS		CURRENT 1-2 KTS	CURRENT > 2 KTS
(Combined Seas)	< 3 FT	3-5 FT	> 5 FT
HUNT	WIND < 20 KTS	WIND 20-30 KTS	WIND > 30 KTS
(Combined Seas)	< 3 FT	3-5 FT	> 5 FT
SWEEP	< 3 FT	3-6 FT	> 6 FT

Compiled from a combination of manuals, including FM 34-81, FM 34-81-1, Service manuals, and various equipment technical orders to present a general picture of METOC impacts to operations. This list is not all-inclusive, nor is it intended to restrict SMOs and/or JMOs to these limitations. SMOs and JMOs should feel free to use this list as a baseline, expanding or changing it as needed to suit the forces and limitations of their various commands.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B

METOC Forecast Element	Green Criteria	Yellow Criteria	Red Criteria	Mission Areas	Notes	Addressed in Literature Search	Addressed on FNMOC Metric site	Summary and Comments
WIND		5-10 kts	< 5 kts or > 19 kts, or wind from target area	Smoke		Martin (2002) highlighted importance for ships at sea. Cantu (2001) found 16% of a set of aircraft mishaps related to wind. Klein (2005) focused on >50 kt wind predictability, at timescales of 5+ days. Current ensemble forecast verification metrics (see FNMOC website) may be sufficient for the specific question of sorties, but perhaps need ensemble site specific verification for TCCOR recommendations.	Current model deterministic and probabilistic metrics exist	These thresholds show that various commanders desire wind forecasts to an accuracy of about 0.5 m/s or 1 kts from 0-60 kts, whereas historically operational forecasters typically consider a wind forecast verified if it is within +/- 5 kts. Therefore, for example, a field forecast of 15 kts might verify as correct at ≤ 13 or ≥ 18 kts, spanning the red/yellow/green criteria for paradrop. Of note, since circa 2014 the Navy TAF verification program (adapted from the National Weather Service) typically uses +/- 2 kts for wind verification (with wind forecasts given in 5 kt increments; Fleet Weather Center San Diego, personal communication, 2017). This may be a case where a customer could be asked to set probabilistic thresholds. Typical wind probability products show probability of exceedance of a given speed, in increments of 5
		< 10 KTS	> 30 kts, or wind calm, or wind from target area	NBC				
	Crosswind < 10 kts	Crosswind 10-15 kts	Crosswind > 15 kts	UAS operations				
	< 10 kts	10-34 kts	> 34 kts	Bridging				
	< 13 kts	13-18 kts	> 18 kts	Paradrop				
	< 20 kts	20-30 kts	> 30 kts	Naval minehunting				
	< 20 kts	20-35 kts	> 35 kts	Sea port operations				
	< 25 kts	25-35 kts	> 35 kts	CAS or Deep attack/Mine Warfare Aviation	generic airframe/Navy MIW aircraft			
	< 25 kts	25-50 kts	> 40 kts	Helicopter (attack)	Discrepancy between yellow and red is as written			
	< 30 kts	30-35 kts	> 35 kts	Artillery				

			> 60 kts	Antisurface warfare	For both ASUW and amphib warfare, seas restrictions inherently contain wind restrictions			kts, usually staring at 20 kts, so in this case a probability of exceedance for something like 13 kts might be more appropriate. It would also be of benefit if the criteria could be calibrated to forecast length, e.g., perhaps 40% chance of > 13 kt winds would be in the "yellow" range for a 72 hour forecast, but in a "green" or "red" range for a 12 hour forecast.
VISIBILITY	> 1600 m	800-1600 m	< 800 m	Helicopter (lift)		Cantu (2001) found visibility to be a factor in over half of a set of aviation mishaps. Darnell (2006) noted surface visibility as one of three elements causing the most negative impacts to missions.	For some stations, broken up by aerosol and hydrometeor impacts	Terminal Aerodrome Forecasts report visibility in increments of 100 m (see for example NAVMETOCCOMINST 3143.1H). However, for forecast validation and verification, MOS visibility categorical levels are used, that have at best a resolution of 1/2 nm (Fleet Weather Center San Diego, personal communication, 2017). Model Output Statistics (MOS) are typically very useful to aviation forecasters, but that is typically only in CONUS. This is consistently considered a hard forecast element. May wish to
	> 2000 m	1000-2000 m	< 1000 m	Armor gun sights				
	> 2500 m	1000-2500 m	< 1000 m	Copperhead				
	> 3000 m	1000-3000 m	< 1000 m	Ground RECCE and Artillery				
	> 3000 m	2000-3000 m	< 2000 m	TOW Missile				
	> 3200 m	1600-3200 m	< 1600 m	CAS or Deep attack				
	> 4000 m	1000-4000 m	< 1000 m	Helicopter (attack)				
	> 4800 m	900-4800 m	< 900 m	Air ports				

	> 4800 m	1000-4800 m	< 1000 m	Amphibious warfare				consider a more integrated aerosol and hydrometeor combined product, vice more separate forecast techniques (i.e., aerosol surface visibility is available on WxMap).
	> 4800 m	3200 - 4800 m	< 3200 m	Predator				
	> 5000 m	3000-5000 m	< 3000 m	Hellfire missile				
	> 5000 m	< 5000 m		Air defense				
	> 7000 m	500-7000 m	< 500 m	Lock on before launch				
	> 7000 m	1700-7000 m	< 1700 m	Lock on after launch				
	> 8000 m	3200-8000 m	< 3200 m	CAS (planning purposes, Army)				
	> 8000 m	4800-8000 m	< 4800 m	Aerial RECON and Airlift				
CEILING	> 500 ft	300-500 ft	< 300 ft	Helicopter (lift)		Darnell (2006) noted cloud ceilings as one of three elements causing the most negative impacts to missions.	Ensemble verification of cloud cover was noted. No easily tracked vertical verifications (except of humidity at standard levels) noted. Possibly base on RH threshold	METEOGRAMS available on demand; possible area where NRL MMD might want to consider some aggregate metric of the elements that go into a ceiling forecast on some geographic scale. Typically, TAFs forecast changes in ceiling down to 100 ft increments.
	> 1000 ft	500-1000 ft	< 500 ft	Airlift				
	> 1500 ft	200-1500 ft	< 200 ft	Air ports				
	> 1500 ft	600-1500 ft	< 600 ft	Artillery				
	> 1700 ft	800-1700 ft	< 800 ft	Lock on after launch				
	> 1900 ft	400-1900 ft	< 400 ft	Lock on before launch				
	> 2000 ft	800-2000 ft	< 800 ft	Hellfire missile and Predator				
	> 2000 ft	1000-2000 ft	< 1000 ft	CAS (planning purposes, Army)				
	> 2600 ft	1100-2600 ft	< 1100 ft	Helicopter (attack)				
	> 3000 ft	1000-3000 ft	< 1000 ft	Copperhead				

	> 3500 ft	1000-3500 ft	< 1000 ft	Flight ops				
	> 5000 ft	300-5000 ft	< 300 ft	Amphibious warfare				
	> 5000 ft	2500-5000 ft	< 2500 ft	Air defense				
CLOUD COVER	< 2/8	2/8 - 4/8	> 4/8	Aerial Recon		Darnell (2006) noted cloud layers as one of three elements causing the most negative impacts to missions.	See ceiling comments	See ceiling comments
	CLR-SCT	BKN	OVC	EO Support				
	No clouds at flight level	SCT-BKN clouds at flight level	OVC at flight level	Aerial refueling				
ICING	None	LGT or MDT	SVR	Helicopter (lift)		Cantu (2001) found icing to be a factor in 3% of a set of aviation mishaps.	See ceiling comments	See ceiling comments; noted cloud ice an option for some models via METCAST
TURBULENCE	None	LGT or MDT	SVR	Helicopter (lift)		Cantu (2001) found turbulence to be a factor in 2% of a set of aviation mishaps.	No, though CAT product available in METCAST - very difficult to verify except via PIREPs	Navy forecasters typically begin with USAF forecasts products and modify as necessary
		LGT or MDT	SVR	CAS or Deep attack (no specific airframe)				
TEMPERATURE	> 20F and < 85F	85 to 95F or -15 to 20F	> 95F or < -15F	Personnel			Relatively robust at synoptic and mesoscale level, at 2m height and a few levels	Duration of temperature may be important as well, possibly an area for a probabilistic threshold at longer lead times.
		85 to 120F	< 32F or >120F	SIGINT				
			> 90F	Helicopter (attack)				
			> 103F	Navy ASUW				
			> 120F	Air defense				
			< 20F or > 125F	Artillery				
PRECIPITATION	NONE	LGT	MDT/HVY	Paradrop and NBC Operations		Cantu (2001) found precipitation to be a factor in 12%	Multiple verification scores available; time and space	Probabilistic criteria may be useful here based on QPF and re-negotiation with operational commanders
	NONE	LGT	MDT	Personnel				

	NONE	MODERATE	HEAVY	Helicopter (attack), Smoke, Copperhead, Artillery, Trafficabilty		of a set of aviation mishaps.	resolution may not be sufficient to meet forecaster desire or commander need	
	No/LGT	MDT	HVY	Over the horizon targeting (OTHT)				
		LGT Freezing precip	Freezing precip that closes runway	Airlift				
			HVY	EO support				
THUNDERSTORMS	No	Few	Scattered	Helicopter (attack), Aerial refueling		Cantu (2001) found thunderstorms to be a factor in 3% of a set of aviation mishaps.	No model derived CAPE or LI (for some set of standard locations) noted as verified	Difficult to provide verifiable products away from CONUS or in open sea
DENSITY ALTITUDE	< 4000 ft	4000-6900 ft	> 6900 ft	Paradrop		Cantu (2001) found density altitude to be a factor in 5% of a set of aviation mishaps.	Not noted	
ATMOSPHERIC STABILITY	Stable	Neutral	Unstable	NBC			Not noted	
ABSOLUTE HUMIDITY	< 14 g/m3	14-18 g/m3	> 18 g/m3	EO Support			Not noted	Relative humidity more typical for typical Navy forecaster
TRANSMITTANCE	> 0.4	0.2-0.4	< 0.2	EO Support			Not noted	
ILLUMINATION		Moonrise/Moonset	No Moon	EO Support			Not noted	
			> 0.0001 ft candles	Landing craft				
MODIFIED SURF INDEX	< 8	8 to 10	> 10	Landing craft			Not noted	
WAVE BREAKER HEIGHTS	< 5 ft	5-12 ft	> 12 ft	Landing craft			Not noted	

WAVE PERIOD	> 7 s	6-7 s	< 6 s	Landing craft			Noted	
SEAS	< 3 ft	3-5 ft	> 5 ft	Divers, Mine hunting		Martin (2002) highlighted importance for ships at sea. Cantu (2001) found 5% of a set of aircraft mishaps related to combined sea state. Klein (2005) discussed probabilities and uncertainty, especially for the probability of seas > 12 ft both by area and arrival time.	Mostly for significant and peak wave height	Seas forecasts (i.e., significant wave height and/or combined wind and swell) desired to have an accuracy of ~1 ft. Extensive efforts have been made by NRL MMD and others, including development of a Wavewatch III model field derived from the human forecaster generated tropical cyclone track.
	< 3 ft	3-6 ft	> 6 ft	Mine sweeping				
	< 6 ft	6-8 ft	> 8 ft	ASUW				
CURRENT		1-2 kts	> 2 kts	Divers			Not noted	Expect other operations to have varied thresholds here as well

APPENDIX C

Survey on METOC Product Effectiveness Conducted at Naval Postgraduate School:

1. Are you familiar with Joint Publications 3-59 Meteorological and Oceanographic Operations, Appendix G (METOC impacts to selected operations)? [*Respondents could select "Yes" or "No" from a drop-down menu]*
2. If you are not familiar with JP 3-59, please review the attachment contained in the email [*invitation to survey*]. Do you agree with the parameters listed for your job specialty? If no, what parameters do you not agree with? [*comment section available*]
3. Did the forecast you received for your mission include these parameters? [*Respondents could select "Yes" or "No" from a drop-down menu]*
4. Within the past five years, did your job include the usage of METOC products? [R*espondents could select "Yes" or "No" from a drop-down menu]*
5. Based on the previous question, in the text box provided specify the level of command the job was performed at. [*Drop-down menu with choices of O-3, O-5, O-6, or O-8 level command*]
6. Based on previous questions, please answer the following: [*Respondents could select "Strongly Agree", "Agree", "Neutral", "Disagree", or "Strongly Disagree" from a drop-down menu]*
    a. The METOC products were useful.
    b. METOC products and forecasts are accurate.
    c. The forecasts I received accurately predicted the weather I encountered.
    d. Forecast accuracy impacts my operations.
    e. Operational planning depends upon available METOC products.
    f. The lead time of the forecasts was sufficient.
    g. A longer lead time for forecasting will benefit operations.
    h. Longer forecasting lead time, but with greater uncertainty, is acceptable.
7. What was the lead time for the forecast you received? [*Respondents could select "less than 6 hours", "between 6 and 12 hours", "between 12 hours and 24 hours", or "greater than 24 hours" from a drop-down menu]*
8. Based on the previous answer, was the lead time sufficient? [*Respondents could select "Yes" or "No" from a drop-down menu]*
9. If a forecast was available at a longer lead time, would you use it? How far in advance? [*Respondents could select "Yes" or "No" from a drop-down menu, and there was a space provided for comments]*
10. Would you use a forecast with a longer lead if the information carried more uncertainty? [R*espondents could select "Yes" or "No" from a drop-down menu]*

11. Based on the previous question, how uncertain can a forecast be before it is no longer useful? [*Respondents could select “Extremely Inaccurate”, “Very Inaccurate”, “Moderately Inaccurate”, “Slightly Inaccurate”, or “Inaccurate” from a drop-down menu]*
12. What is the latest you can receive accurate METOC products before it significantly disrupts planned operations? Please respond in terms of days. [*Respondents could select from 1-30 days from a drop-down menu]*
13. What was the most useful METOC product you have received for your operation? *[A space was provided for comments]*
14. What was the least useful METOC product you have received for your operation? *[A space was provided for comments]*
15. Do you have any additional thoughts, anecdotes, complaints or other items in regards to METOC support? *[A space was provided for comments]*
16. Have you had formal Meteorology and/or Oceanography (METOC) training from the military or separate education, to include space weather? If yes, please indicate which one. *[Check-boxes were provided for Meteorology, Oceanography, and Space Weather]*
17. *Additional metadata questions were included on branch of service, highest rank, years of military service, whether the respondent had ever held command, and if so, at what level did they command*

LIST OF REFERENCES

Butler, M., 2005. The Impacts of Weather Forecasts on Military Operations: A System for Conducting Quantitative Near-Real Time Analyses. M.S. thesis, Naval Postgraduate School Monterey CA, 101 pp.

Callahan, J., 2006. Metrics of METOC forecast performance and operational impacts on carrier strike operations. M.S. thesis, Naval Postgraduate School Monterey CA, 116 pp.

Cantu, R.A., 2001. The Role of Weather in Class A Naval Aviation Mishaps FY90-98. M.S. thesis, Naval Postgraduate School Monterey CA, 103 pp.

Cunningham, J.G., 2006. Applying ensemble prediction systems to Department of Defense Operations. M.S. thesis, Naval Postgraduate School Monterey CA, 149 pp.

Commander, Naval Meteorology and Oceanography Command, 2018: Commander Naval Meteorology and Oceanography Command Strategic Plan, 10 pp, accessed at https://usff.navy.deps.mil/sites/cnmoc/instructions/default.aspx.

Darnell, K., 2006. Analysis of weather forecast impacts on United States Air Force combat operations. M.S. thesis, Naval Postgraduate School Monterey CA, 121 pp.

Heacox, Nancy J., and Coauthors, 2004: Real-time online communications: 'Chat' use in Navy operations. Space and Naval Warfare Systems Command San Diego CA, 12 pp, https://apps.dtic.mil/dtic/tr/fulltext/u2/a465828.pdf.

Joint Chiefs of Staff, 1999: Joint Pub 3-59, Joint Doctrine, Tactics, Techniques, and Procedures for Meteorological and Oceanographic Operations, 111 pp, accessed at www.dtic.mil.

Katz R., and A. Murphy, editors, 1997: *Economic Value of Weather and Climate Forecasts.* Cambridge University Press, 222 pp.

Klein, P., 2005. The four Florida hurricanes of 2004 and their impact on the Fleet (No. NRL/MR/7005--05-8893). Naval Research Lab Washington DC, 30 pp, accessed at www.dtic.mil.

Lind, A., Dershowitz, A. and Bussolari, S.., 1994: The influence of Data Link-provided graphical weather on pilot decision-making (No. MIT-ATC-215). Massachusetts Inst of Tech Lexington Lincoln Lab, 84 pp, accessed at www.dtic.mil.

Martin, B.S., 2002. METOC and Naval Afloat Operations: Risk Management, Safety, & Readiness. M.S. thesis, Naval Postgraduate School Monterey CA, 96 pp.

Nadeau, L, 2017: Developing Societal Outcome Indicators for NWS' Weather Ready Nation Program. *12th Symposium on Societal Applications: Policy, Research and Practice,* Seattle, WA, Amer. Meteor. Soc., https://ams.confex.com/ams/97Annual/webprogram/Paper302661.html.

National Weather Service, 2011: NOAA'S NATIONAL WEATHER SERVICE STRATEGIC PLAN: Building a Weather-Ready Nation, 44 pp, https://www.weather.gov/media/wrn/strategic_plan.pdf.

Palmer, J., 2010: Incorporating ensemble-based probabilistic forecasts into a campaign simulation in the Weather Impact Assessment Tool (WIAT). M.S. thesis, Naval Postgraduate School Monterey CA, 137 pp.

Robinson, Michael, Joseph C. Venuti, and Starr McGettigan, 2011:. Use of a High-Resolution Deterministic Weather Forecast for Strategic Air Traffic Management Decision Support. Massachusetts Inst of Tech Lexington Lincoln Lab, 26 pp, accessed at www.dtic.mil.

Szczes, J. R., 2008: Communicating optimized decision input from stochastic turbulence forecasts. M.S. thesis, Naval Postgraduate School Monterey CA, 159 pp.

United States Forces Japan, 2015: Tropical Cyclone Conditions of Readiness (TCCOR) (USFJ Instruction 15-4001), 30 pp, https://www.usfj.mil/Portals/80/Documents/Instructions/15-4001%20(USFJI).pdf.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Research Sponsored Programs Office, Code 41
   Naval Postgraduate School
   Monterey, CA 93943

4. Naval Research Laboratory Marine Meteorology Division
   Monterey, California

5. U.S. Army Training and Doctrine Command (TRADOC) Analysis Center
   (TRAC) Monterey
   Monterey, California