COMBAT-AFRL/711: A Tool for Integrating Human Factors into Wargames

Megan B. Morris, Christopher A. Stevens	Bella Z. Veksler	Emma Robin, Kevin Contreras
Air Force Research Laboratory	TiER1 Performance Solutions	Booz Allen Hamilton
Wright-Patterson AFB, OH	Covington, KY	Arlington, VA
megan.morris.3@us.af.mil, christopher.stevens.28@us.af.mil	b.veksler@tier1performance.com	Robin_Emma@bah.com, Contreras_Kevin@bah.com

ABSTRACT

Human factors, such as fatigue and injury, are often represented in wargames at a low level of fidelity. Although this approach enables performance impacts to be considered in planning and adjudication, it limits the generalizability of insights. Modeling and simulation tools for human factors effects are available and improving in resolution. However, an infrastructure does not exist for integrating these tools into wargames or with digital tools commonly used for wargames. To address this gap, a pipeline was created to integrate two models of human performance with an existing digital gameboard tool, developed in support of wargames, Combat Operations Mission and Basing Analysis Tool (COMBAT). One model represents pilot alertness as a function of circadian rhythms, sleep history, and hours awake. Over a 24-hour period, the model can identify periods of low risk (green), moderate risk (yellow) and high risk (red). The second model represents effects of a ground-based laser dazzle weapon on visibility during takeoff and landing. When integrated, the resulting tool, 'COMBAT-AFRL/711', shows planned sorties that are affected by risk associated when either of these factors are flagged. Players and adjudicators may choose to remove flagged sorties based on risk posture or scenario rules. Users may specify sortie schedules and base location of laser dazzle risks. Once the scenario is run, analytics are generated showing users the mission impacts of the human factor risks in the scenario, including canceled sorties and incomplete mission objectives. The tool was applied to data from an unclassified wargame scenario. The analysis of the unclassified scenario revealed significant immediate and cascading effects of canceled sorties on mission objectives. This is important because it highlights ways in which fatigue risks may have unforeseen, downstream operational consequences. The tool enables integration of high-fidelity human factors effects into large mission-level wargames.

ABOUT THE AUTHORS

Dr. Megan Morris is a research psychologist and the Digital Models of Cognition Core Research Area Lead at the Air Force Research Laboratory's Human Effectiveness Directorate. She received her Ph.D. in Human Factors and Industrial/Organizational Psychology from Wright State University in 2014.

Dr. Christopher Stevens is a senior research psychologist and the Learning and Operational Training Core Research Area Lead at the Air Force Research Laboratory's Human Effectiveness Directorate. He received his Ph.D. in Psychology from Penn State University in 2014.

Dr. Bella Veksler is a cognitive science researcher with TiER1 Performance Solutions at the Air Force Research Laboratory. She received her Ph.D. in Cognitive Science from Rensselaer Polytechnic Institute in 2011.

Ms. Emma Robin is a Data Scientist within Booz Allen Hamilton's Global Defense Sector. She holds a BA in Mathematics from Wesleyan University and is concluding an MS in Business Analytics at the Georgia Institute of Technology.

Mr. Kevin Contreras is a Chief Scientist within Booz Allen Hamilton's defense analytics business where he leads the development of decision-support solutions for Air Force and DoD customers. He holds a BS from Georgetown University and an MS from Johns Hopkins University.

COMBAT-AFRL/711: A Tool for Integrating Human Factors into Wargames

Megan B. Morris, Christopher A. Stevens
Air Force Research Laboratory
Wright-Patterson AFB, OH
megan.morris.3@us.af.mil, christopher.stevens.28@us.af.mil

Bella Z. Veksler
TiER1 Performance Solutions
Covington, KY
b.veksler@tier1performance.com

Emma Robin, Kevin Contreras
Booz Allen Hamilton
Arlington, VA
Robin_Emma@bah.com, Contreras_Kevin@bah.com

INTRODUCTION AND BACKGROUND

Wargaming is a critical tool in the Warfighter arsenal to help train and plan wartime strategy and maneuvers. Although wargaming has existed for centuries, only recently have digital simulation capabilities been integrated into scenarios (Appleget, 2021; Caffrey, 2000). These capabilities have greatly advanced the ability of wargames to be more realistic and effective in terms of simulating logistics and material resources. However, human resource and performance fidelity is lacking (Appleget, 2021), decreasing the overall potential of these tools. In wargames, to the extent they are considered at all, humans are typically modeled as optimized assets that are easily regeneratable - far from the reality of operational missions. Recent efforts have been made to take this limitation into account and increase human factors fidelity in wargaming (e.g., Morris et al., 2022). The current effort focuses on developing a digital gameboard that enables wargame players and adjudicators to plan, simulate, and adjudicate missions and sorties while accounting for risks associated with human factors. Specifically, fatigue and directed energy modeling capabilities are integrated into the Combat Operations Mission and Basing Analysis Tool (COMBAT), creating a new implementation of the software, COMBAT-AFRL/711.

COMBAT Software

COMBAT is a digital gameboard framework, developed by Booz Allen Hamilton, that integrates multi-domain models and disparate data types into a single environment to provide end-to-end simulation for wargame move analysis. The software was developed using R (R Core Team, 2022) with the *RShiny* package (Chang et al., 2023) for the user interface, and the framework is based on data fields and scenarios commonly found within Air Force Futures Title 10 Wargames (Bradford, 2022). These wargames include a digitized modeling suite that includes several mission-specific tools, such as base defense and aerial refueling, among others. However, these tools do not take into consideration human factor elements, limiting their fidelity for assessing mission and campaign-level operations.

Fatigue Modeling Capabilities

A pervasive issue in military communities is cognitive fatigue resulting from sleep restriction/deprivation, circadian desynchrony, time-on-task, and high levels of cognitive load (Caldwell, 2005; Miller et al., 2011). Fatigue is associated with reduced performance and costly mishaps across Department of Defense (DoD) services (e.g., Gaines et al., 2020; Kelley et al., 2018; United States Navy, 2017). Wartime is especially fatiguing given the extreme demands and operation tempo needed to compete with adversaries. As such, it is critical to represent fatigue in wargaming scenarios and how it affects the performance effectiveness and readiness of Warfighters. An initial effort was conducted by the Air Force Research Laboratory (AFRL) to model fatigue effects associated with data from a wargaming logistics software (Morris et al., 2022). Specifically, an application was developed in R and *RShiny* that ingested Air Tasking Order (ATO) inputs from Frontier Technology Inc.'s Integrated Sustainment Wargaming and Analysis Toolkit (iSWAT) and produced fatigue estimates for aircrew and maintenance personnel based on a biomathematical fatigue model and algorithms for realistic sleep schedules. The Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE; Hursh et al., 2004) biomathematical model produces a performance effectiveness value from 0 to 100%, with typically 77.5% and above reflecting low fatigue and good performance, 70% to 77.5% reflecting medium fatigue where individuals need to think about mitigation strategies, and 70% and below reflecting high fatigue where mitigation is needed immediately. Performance effectiveness at 70% has been found to be equivalent to a .08%

blood alcohol concentration (Hursh et al., 2006) in respect to effects on some cognitive processes. Along with this biomathematical model, algorithms were developed with input from subject matter experts to develop generalized sleep schedules for aircrews and maintainers. Using a mock-up wargaming scenario with optimal and realistic sleep scenarios, the effort found that certain aircrews were experiencing high levels of fatigue, creating safety risks and likely degrading performance. These effects were not captured in the wargame scenarios without the fatigue modeling and have important implications as to the performance to expect from these crews, their readiness throughout the scenario, and cascading effects on related missions. The modeling capabilities of this initial effort were integrated into the COMBAT software and will be further discussed in the Method section.

Laser Dazzle Modeling Capabilities

Directed energy weapons are a critical concern for the military as adversaries have harnessed these technologies to degrade and defeat operations. Directed energy weapons are electromagnetic systems that convert electrical or chemical energy to focused, radiated energy (Hoffman, 2022). This focused energy is targeted on a range of entities such as electronic devices, vehicles, missiles, and humans. Directed energy's effects on humans is referred to as bioeffects, how energy affects living tissue. Laser dazzle is a common method of disrupting operations, where laser energy is dispersed across an individual's retina, causing temporary, or possibly permanent, visual loss (Williamson & McLin, 2018). Given the effects of laser dazzle, it is important to be able to model its bioeffects to inform prediction of effects on human performance and to inform methods of protecting our military personnel. Researchers at AFRL along with collaborators have developed models of laser dazzle effects on human perception, showing effects on specific perceptual processes such as saccadic targeting and feature extraction (e.g., Williamson et al., 2017; see Williamson et al., 2013 and Williamson, 2016 for additional information on models of laser dazzle effects). We integrated one of these models into the COMBAT software and this will be further discussed in the Method section.

METHOD

Human Factors Modeling Integration

In order to develop the new digital gameboard software, the framework and methodology from the original COMBAT software was leveraged for grouping missions, flight paths, tracking engagements, and map visualizations. This framework is based on Future Games 2023 data fields and scenarios. Four mission types were focused on: Offensive Counter-Air (OCA), Defensive Counter-Air (DCA), Fires (where targets are destroyed on the ground from the air), and Aerial Refueling (AR). Aircraft types for these missions included fighters, bombers, unmanned aircraft vehicles (UAVs), and tankers.

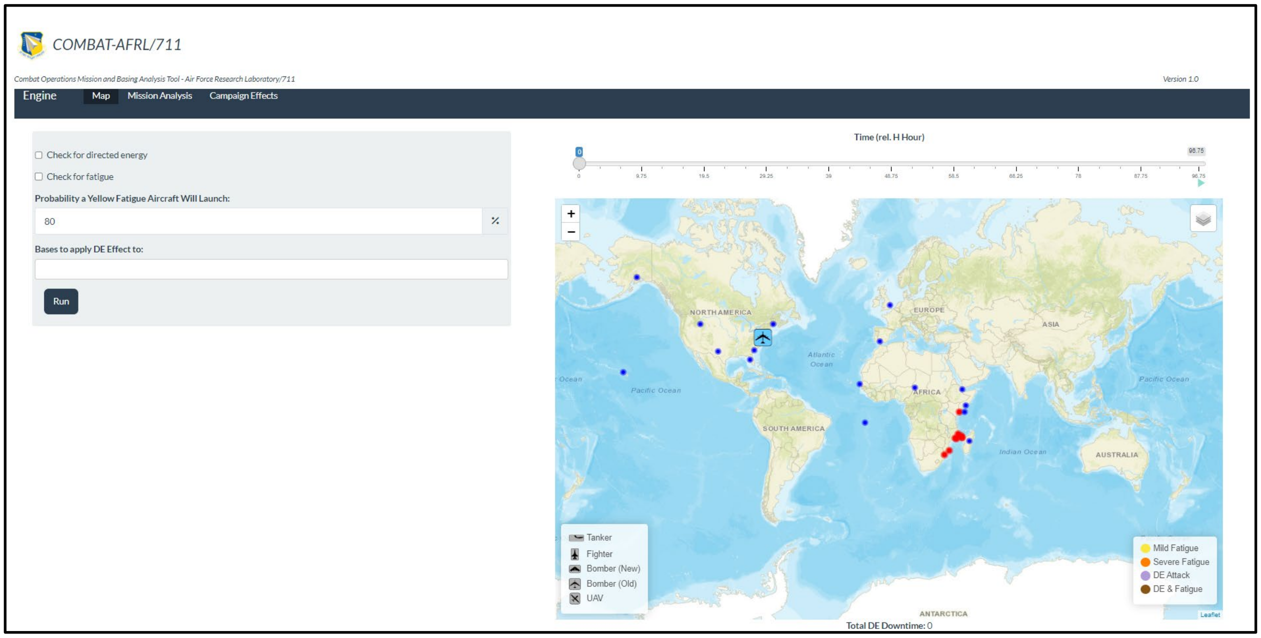
Serendipitously, COMBAT and the fatigue modeling software were both developed using R and *RShiny*, so it was most beneficial to continue using this program (R version 3.6.3) to develop COMBAT-AFRL/711. In order for the underlying models to inform each data source, R package *RSQLite* (Müllert et al., 2023) was also used to embed a relational database management system within the framework that interfaces with R Studio (RStudio Team, 2020).

To model fatigue, the following data was used from simulated sorties: the departure time, arrival time, location (base), type of aircraft, and number of aircraft performing the mission. Fatigue modeling assumptions from previous work (Morris et al., 2022) were used in the current framework. Specifically, a work period comprised the 2.5 hours before a mission, the mission length, and 2 hours after the mission to represent briefing and debriefing time periods. The following were used to categorize fatigue effects on work periods: 1) *No Fatigue*, where pilot performance effectiveness was above 78% and it is assumed that the associated mission launched. 2) *Mild Fatigue*, where pilot performance effectiveness was between 70% and 78%, and the associated mission is flagged and launch status is based on probability and user input. 3) *Severe Fatigue*, where the pilot performance effectiveness was below 70%, and the mission did not launch; however, users can override this through manually changing the launch status. For visualization purposes, the software calculated the maximum amount of fatigue a pilot experienced at any point during their work period and missions that fell under *Mild* or *Severe* fatigue were flagged and displayed. It should be noted that the fatigue model does not account for sorties that exceed 24 hours. These sorties were treated as *Mild* fatigue, as there were very few instances, and it is likely these would include augmented crews that will allow for some crew rest to reduce fatigue.

Directed energy model inputs were integrated as a probabilistic baseline and are based on user input where bases and locations have to be selected for directed energy effects to occur. The following were used to categorize directed energy effects: 1) *DE Impact*, where the pilot experienced directed energy dazzle while landing. 2) *No DE Impact*, where the pilot did not experience a directed energy dazzle while landing. If a pilot experienced a directed energy dazzle, the software then canceled all subsequent missions for that aviation platform at the same base. Note, that the nature of directed energy effects on perception are more severe compared to fatigue – laser dazzle fully incapacitates a pilot; fatigue does not necessarily fully incapacitate a pilot, but rather degrades cognitive processes.

Each modeling capability (i.e., fatigue and directed energy) utilized its own assumptions and thresholds (as mentioned above) for determining mission impact. Once a threshold was crossed for determining that a mission was canceled, it was marked as such in the scenario; however, it is possible for an adjudicator to overrule this cancelation. Model outputs from the new integrated software were tested against the original baseline models to validate output.

An interactive dashboard and map were developed to show campaign results, visualize the mission and model outputs, and house data (see Figure 1). Fatigue and directed energy impacts are shown in both tabular form on the homepage (this is where users can change the launch status of a mission) and within the campaign map.



☐ Check for directed energy

☐ Check for fatigue

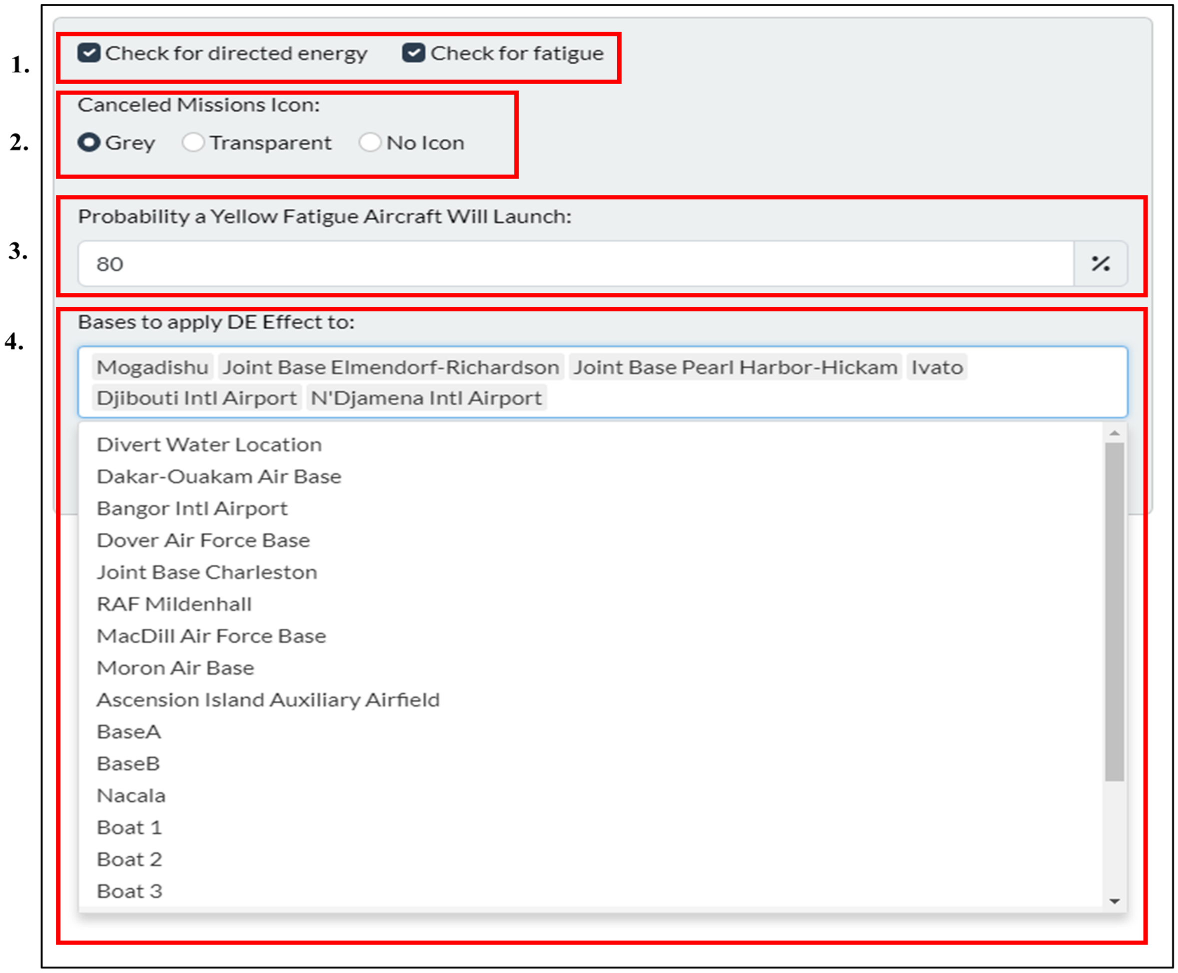
Probability a Yellow Fatigue Aircraft Will Launch: 80 %

Bases to apply DE Effect to:

Total DE Downtime: 0

Figure 1. COMBAT-AFRL/711 homepage.

The homepage has a configuration panel where a user can 1) select which human performance parameters to customize, 2) select the indicator for canceled missions, 3) change the probability that an aircraft with a pilot with *Mild Fatigue* will launch, and 4) select which bases directed energy effects are applied to (see Figure 2).



1. ☑ Check for directed energy ☑ Check for fatigue

2. Canceled Missions Icon:
● Grey ○ Transparent ○ No Icon

3. Probability a Yellow Fatigue Aircraft Will Launch:
80 %

4. Bases to apply DE Effect to:
Mogadishu, Joint Base Elmendorf-Richardson, Joint Base Pearl Harbor-Hickam, Ivato, Djibouti Intl Airport, N'Djamena Intl Airport

Divert Water Location
Dakar-Ouakam Air Base
Bangor Intl Airport
Dover Air Force Base
Joint Base Charleston
RAF Mildenhall
MacDill Air Force Base
Moron Air Base
Ascension Island Auxiliary Airfield
BaseA
BaseB
Nacala
Boat 1
Boat 2
Boat 3

Figure 2. COMBAT-AFRL/711 configuration panel.

The homepage also has a map panel (see Figure 3) to visualize bases and aircraft movement throughout the campaign, and a mission analysis interactive dashboard focused on the human factors campaign outputs. Within the map, base indicators (rings around bases) and mission indicators (coloring of aircraft) were used to reflect fatigue and directed energy effects based on the current timestamp of the mission. Base indicators reflect the type and volume of missions impacted by fatigue, directed energy, or both. Specifically, yellow is *Mild Fatigue*, orange is *Severe Fatigue*, purple is a *Directed Energy (DE) effect*, and brown is a *Directed Energy and Fatigue Effect* (see Figure 4).

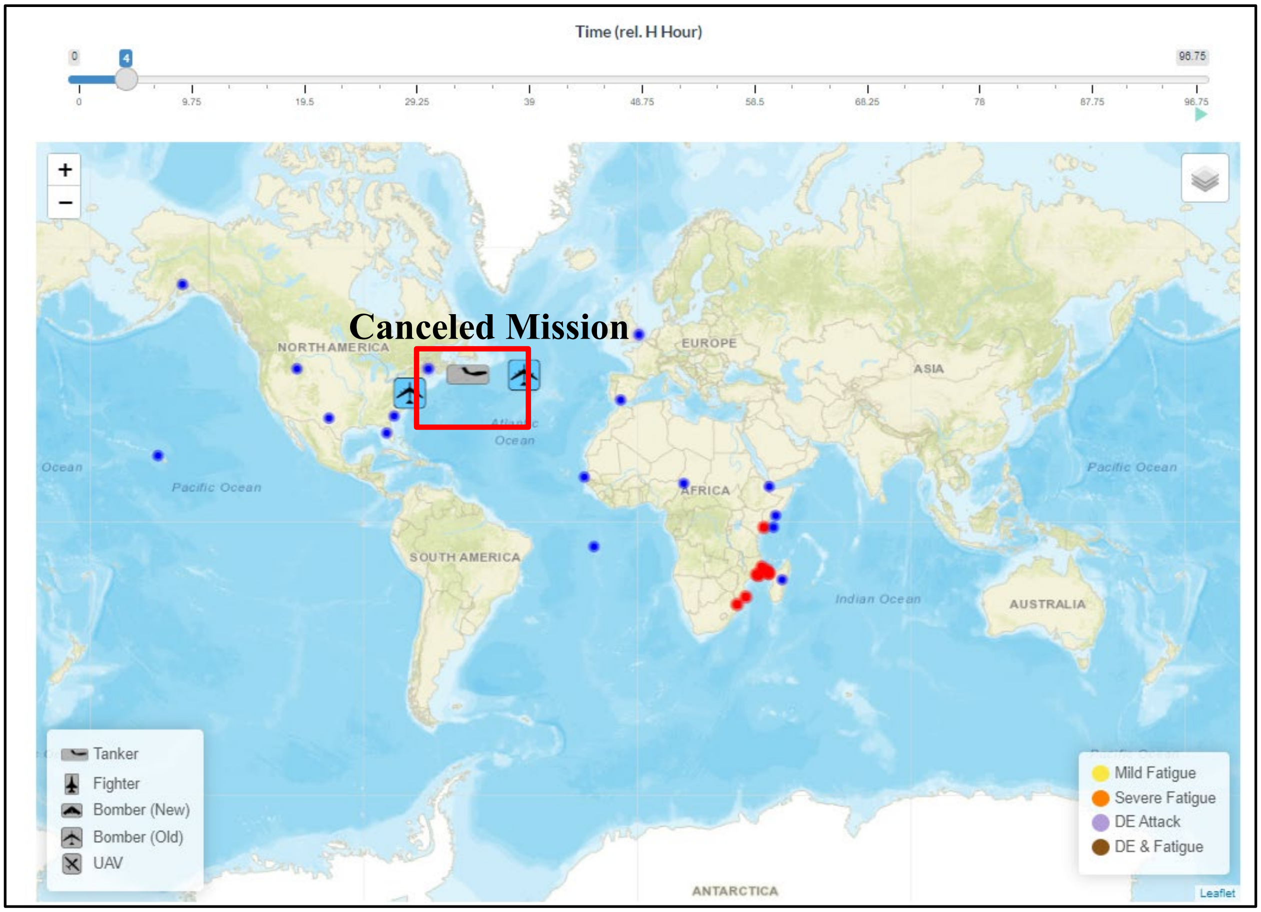


Figure 3. COMBAT-AFRL/711 map panel.

Figure 4. Closeup of base indicators.

The software also includes an interactive scatter plot for mission types, bases, platforms, and outputs, as well as associated graphs showing the overall results from the campaign (see Figure 5).

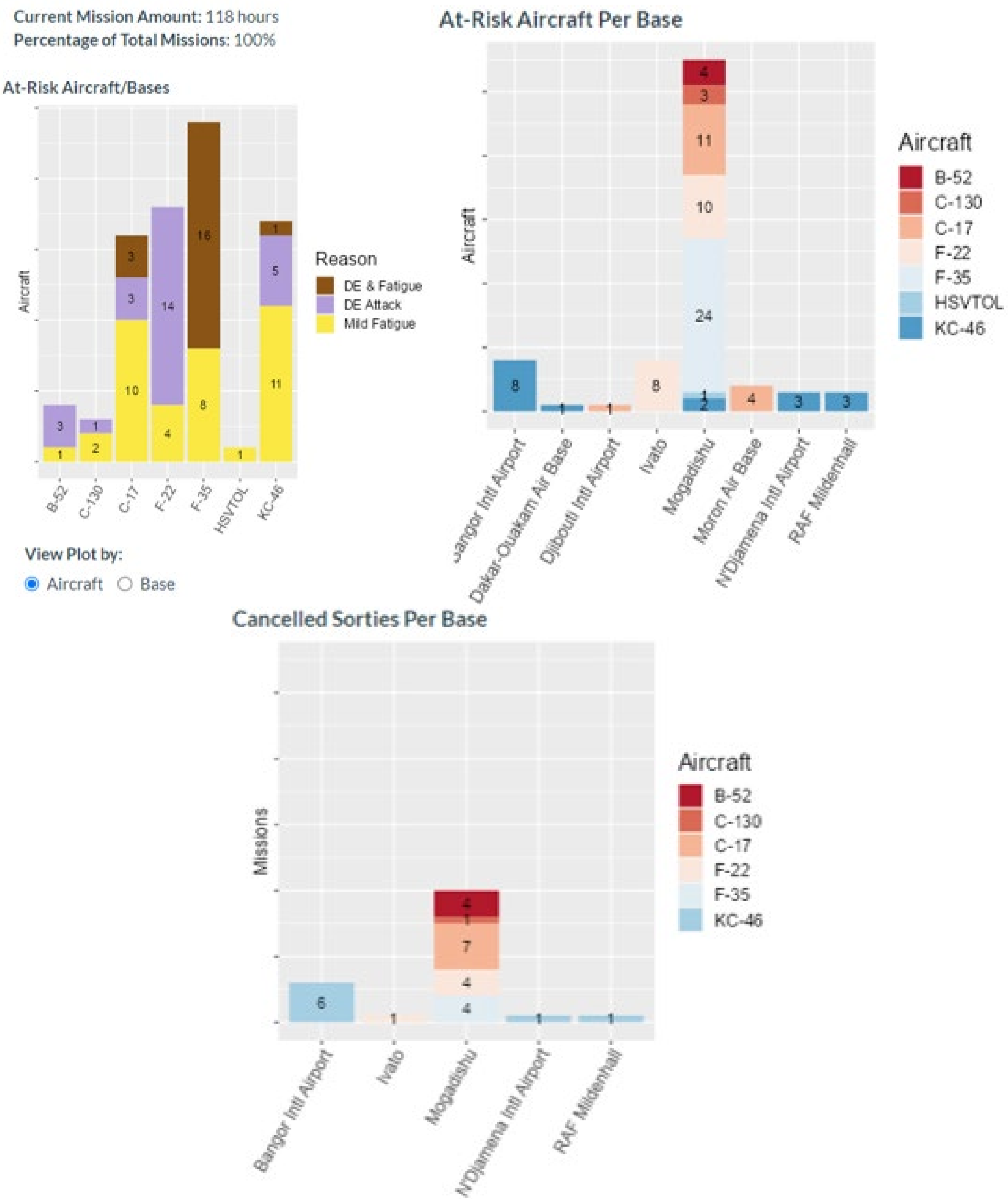


Figure 5. COMBAT-AFRL/711 dashboard graphs.

Wargaming Scenario

To demonstrate a proof-of-concept of this software, a synthetic unclassified wargame scenario was generated to show effects of fatigue and directed energy modeling. The scenario included a fictional conflict with Blue forces located in North America and Europe and Red forces centered in Africa. The scenario included mission types similar to those featured in Future Games 2023 through a mobility treadmill, bomber treadmill, and DCA and Fires missions. Long duration missions had air refueling points along their mission path. The simulated campaign had a Blue main operating base that was defending an island off the coast of Africa, Madagascar, while the Red team tried to cross the channel to capture Madagascar. This scenario was considered to have "pre-adjudicated results" in that all missions launched and executed.

Analysis

To analyze the effect of human factors modeling within the wargaming scenario, a baseline of campaign outcomes from the scenario with no human factors elements was compared to an implementation within the new software. The flags of fatigue and directed energy effects on missions within the homepage and dashboard were used to calculate the percentage of missions impacted and the percentage of missions canceled given these human factors injects. The cascading effects of these impacts were then examined by quantifying the mission types that were impacted by the effects and canceled missions.

RESULTS

The baseline for the wargaming scenario was created to be logistically feasible, but it did not include any human factors. As a result, all missions launched and executed within this baseline implementation. With the human factor modeling integrated into the scenario, 25% of missions were impacted with *Mild Fatigue* and 11% of missions were canceled due to a *DE attack* at the Blue main operating base. The missions impacted by fatigue would result in pilots having degraded general cognitive effectiveness. This would include reductions in reaction time, discrimination, reasoning, and language comprehension (Hursh et al., 2004), increasing the safety risk of these missions and the probability that they will not be successful. With both models injected into the new software, only 65% of missions can launch without any human factor impacts. Examining cascading effects of these impacts, 25.5% of refueling missions were flagged with *Mild Fatigue*, resulting in 9 tanker refueling points impacted (i.e., 27% of tanker missions) and 2 bomber refueling points impacted (i.e., 50% of all bomber missions). Note, that given the nature of fatigue, longer missions with refueling points were more likely to be impacted by fatigue than other shorter missions due to the nature of the fatigue modeling. *DE attack* cascading effects resulted in an additional 15% of missions canceled, including 50% of bomber refueling missions being canceled. Overall, both fatigue and directed energy cascading effects resulted in 75% of bomber missions impacted and 25% of fires missions impacted – suggesting far reaching impacts of fatigue and directed energy attacks in the wargaming scenario. Note that adjudicators decide if these impacted missions are canceled or not, so impacts can have even more critical effects on the scenario if adjudicators decide to cancel these missions based on scenario factors evolving during the wargame.

DISCUSSION

The current effort is a proof-of-concept that shows one can increase human factors fidelity within wargaming tools and that this integration has impactful effects on decision-making with these tools. Specifically, the effort showed that a significant portion of missions are likely to have crews who are experiencing fatigue. This will likely degrade performance and decision-making with these crews and affect the overall mission performance. These degradations also had cascading effects on additional missions connected to these missions, resulting in increased safety risk. Although a small portion of missions were canceled due to DE attacks, one must take into consideration the cascading effects of these canceled missions as shown in the current effort with impacts on a variety of missions such a refueling and fires. The software gives a theory-driven mechanism to inject these effects into a plausible scenario and can be used as a training device to help participants in wargames realize the impact of human factors and incorporate these likely outcomes into their decision-making.

Limitations

The initial fatigue model integrated into the software does not account for missions that exceed 24 hours or that start before the campaign timeline. In the next iteration of the software, fatigue modeling will be extended beyond 24 hours for missions, as well as those that start before the campaign, as this will be crucial to reflect long-duration missions that may be likely in the future with potential conflict in the Indo-Pacific region. The directed energy model outputs the time the directed energy attack occurred, effecting the pilot at the specific base for a set amount of time. This led to subsequent missions at the base with that same aircraft being canceled. However, it is possible other pilots might be available to pilot the subsequent missions. In future iterations of the software, one would ideally track missions by tail number to specify which missions will not launch. The two modeling tools at present do not take tasks (e.g., refueling vs fires) into consideration, but output effects on general performance. Future work could explore developing predictions for specific mission tasks and how they might be differentially affected by the factor (i.e., fatigue or laser dazzle).

Future Directions

Integrating human factors deeper into wargames through analytically sound methods for data capture, modeling, simulation, and analytics, including the potential of live gameplay for training and research purposes, will be critical to improving decision-making and strategy within these games. In the next iteration of the software, as mentioned above, fatigue modeling will be extended. Related to this, it would be useful to integrate overall readiness modeling, which will include elements of fatigue modeling to inform cognitive readiness of crews. In addition, this readiness modeling will focus on the ability to include personnel factors such as specialty codes and prior training to inform readiness for missions. In addition to modeling advancements, we will increase the usability of the software, focusing on user interface design and associated graphics.

ACKNOWLEDGEMENTS

The opinions expressed herein are solely those of the authors and do not represent the official positions of the United States Government, the U.S. Department of Defense, the U.S. Air Force, or any of their subsidiaries or employees. Distribution A: Approved for Public Release. Case number AFRL-2024-3536.

REFERENCES

Appleget, J. (2021). An introduction to wargaming and modeling and simulation. In C. Turnitsa, C. Blais, & A. Tolk (Eds.), *Simulation and Wargaming* (pp. 1-22). https://doi.org/10.1177/15485129211073126

Bradford, B., & Gellert, F. (2022). Caveat milites: Soldiers, beware – Learning experiences from past Army experiments. https://warroom.armywarcollege.edu/articles/caveat-milites/

Caffrey, M. (2000). Toward a history-based doctrine for wargaming. *Aerospace Power Journal*, *14*(3), 33-56. https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Chronicles/caffrey.pdf

Caldwell, J. A. (2005). Fatigue in aviation. *Travel Medicine and Infectious Disease*, *3*(2), 85–96. https://doi.org/10.1016/j.tmaid.2004.07.008

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2024). *shiny: Web Application Framework for R*. R package version 1.8.0, https://github.com/rstudio/shiny, https://shiny.posit.co/.

Gaines, A. R., Morris, M. B., & Gunzelmann, G. (2020). Fatigue-related aviation mishaps. *Aerospace Medicine and Human Performance*, *91*(5), 440-447. https://doi.org/10.3357/AMHP.5515.2020

Hoffman, R. (2022). Directed Energy Weapons (DEW) High Power Microwave (HPM) 6.1 Programs FY22 Annual Report. (DCN# 0543-1375-23) Office of Naval Research Science & Technology. https://www.nre.navy.mil/media/document/directed-energy-weapons-dew-high-power-microwave-hpm-61-programs-fy22-annual-report

Hursh, S. R., Raslear, T. G., Kaye, A. S., & Fanzone, J. F. (2006). *Validation and calibration of a fatigue assessment tool for railroad work schedules, summary report* (Report No. DOT/FRA/ORD-06/21). U.S. Department of Transportation, Federal Railroad Administration, Office of Research and Development. Retrieved from https://rosap.ntl.bts.gov/view/dot/62575.

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Storm, W. F., Miller, J. C., & Eddy, D. R. (2004). Fatigue models for applied research in warfighting. *Aviation, Space, and Environmental Medicine*, *75*(3), A44-A53. https://www.ingentaconnect.com/content/asma/asem/2004/00000075/a00103s1/art00005

Kelley, A. M., Feltman, K. A., & Curry, I. P. (2018). A survey of fatigue in Army Aviators. *Aerospace Medicine and Human Performance*, *89*(5), 464-468. https://doi.org/10.3357/AMHP.5044.2018

Miller, N. L., Matsangas, P., & Shattuck, L. G. (2008). Fatigue and its effect on performance in military environments. In P. A. Hancock & Szalma, J. L. (Eds.), *Performance under stress* (pp. 231-249). Burlington, VT: Ashgate Publishing, Ltd.

Morris, M. B., Veksler, B. Z., Noesen, B., Tuttle, J., Carpenter, B., Tran, P., & Gunzelmann, G. (2022). Human fatigue modeling in wargaming simulations. *Proceedings of the 2022 Interservice/Industry Training, Simulation, and Education Conference*, Paper No. 22331. https://www.xcdsystem.com/iitsec/proceedings/index.cfm?Year=2022&AbID=112519&CID=944#View

Müllert, K., Wickham, H., James, D. A., & Falcon, S. (2023). *RSQLite: SQLite Interface for R*. R package version 2.3.7. https://github.com/r-dbi/RSQLite, https://rsqlite.r-dbi.org

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

RStudio Team. (2020). *RStudio: Integrated Development for R.* RStudio, PBC, Boston, MA. http://www.rstudio.com/

United States Navy (2017). *Comprehensive review of recent surface force incidents*. Norfolk, VA. Retrieved from https://www.public.navy.mil/usff/Documents/USFF-Comprehensive-Review-2017.pdf

Williamson, C. A. (2016). Simple computer visualization of laser eye dazzle. *Journal of Laser Applications*, *28*, 012003. https://doi.org/10.2351/1.4932620

Williamson, C., & McLin, L. N. (2018). Determination of a laser eye dazzle safety framework. *Journal of Laser Applications*, *30*(3), 032010. https://doi.org/10.2351/1.5029384

Williamson, C. A., McLin, L. N., Rickman, J. M., Manka, M. A., Garcia, P. V., Kinerk, W. T., & Smith, P. A. (2017). Wavelength and ambient luminance dependence of laser eye dazzle*.* *Applied Optics*, *56*(29), 8135-8147. https://doi.org/10.1364/AO.56.008135

Williamson, C. A., Strachan, E., & Siebert, J. P. (2013, March). Simulating the effects of laser dazzle on human vision. In *International Laser Safety Conference* (Vol. 2013, No. 1, pp. 268-277). Laser Institute of America.