



UNITED STATES ARMY AEROMEDICAL RESEARCH LABORATORY

Progress Toward an Operator State Monitoring System for U.S. Army Aviation

**Katie Feltman, Xiaomin Yue, Jonathan Vogl, J. Andrew Atchley,
Jordayne Wilkins, Kevin Baugher, Jared Basso,
& Amanda Kelley**

Notice

Qualified Requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Fort Belvoir, Virginia 22060. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of Address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

Human Subjects Use Statement

In the conduct of research involving human subjects, the investigator(s) adhered to the policies regarding the protection of human subjects as prescribed by Department of Defense Instruction 3216.02 (Protection of Human Subjects and Adherence to Ethical Standards in DoD-Supported Research) dated 8 November 2011.

IRB Determination and Number

Multiple studies are included in this report. The determinations for each are reported below.

The study titled “Development of a Composite Performance Index for Simulator Performance and Determination of Timing Intervals of Physiological Data” USAARL 2024-009 was reviewed by the USAARL Exempt Determination Official (EDO) and was determined to be exempt human subjects research [32 CFR 219.104(d)(4ii)] as the project used existing deidentified data on 17 July 2024.

The study titled “Survey of Army Aviators: Understanding Performance” USAARL 2025-025 was reviewed by the USAARL Exempt Determination Official (EDO) and was determined to be exempt human subjects research [32 CFR 219.104(d)(4ii)] as the survey did not ask for identifying information on 12 May 2025.

REPORT DOCUMENTATION PAGE				<i>Form Approved</i> <i>OMB No. 0704-0188</i>								
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.												
1. REPORT DATE (DD-MM-YYYY) 25-09-2026		2. REPORT TYPE Technical Report		3. DATES COVERED (From - To) 1 OCT 2023 - 30 SEP 2026								
4. TITLE AND SUBTITLE Progress Toward an Operator State Monitoring System for U.S. Army Aviation				5a. CONTRACT NUMBER								
				5b. GRANT NUMBER								
				5c. PROGRAM ELEMENT NUMBER								
6. AUTHOR(S) Feltman, K. A. ¹ , Yue, X. ^{1,2} , Vogl, J. ¹ , Atchley, A. ^{1,2} , Wilkins, J. ^{1,3} , Baugher, K. ¹ , Basso, J. ¹ , & Kelley, A. ¹				5d. PROJECT NUMBER MO240076								
				5e. TASK NUMBER								
				5f. WORK UNIT NUMBER								
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Aeromedical Research Laboratory P.O. Box 620577 Fort Rucker, AL 36362				8. PERFORMING ORGANIZATION REPORT NUMBER USAARL-TECH-TR--2026-42								
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Medical Research and Development Command Military Operational Medicine Research Program 504 Scott Street Fort Detrick, MD 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S) USAMRDC MOMRP								
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)								
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release: distribution is unlimited.												
13. SUPPLEMENTARY NOTES ¹ U.S. Army Aeromedical Research Laboratory, ² Chickasaw Nation Industries, ³ Oak Ridge Institute for Science and Education												
14. ABSTRACT The project covered within this report addressed a number of areas related to the development of an operator state monitoring system for U.S. Army aviation. These included a survey of current aviators, review of research studies, retrospective analysis, implementation of data synchronization techniques, collection of physiological data in operational scenarios, and using physiological data to predict flight-related performance outcomes. From these efforts we were able to confirm prevalent occupational stressors impacting performance include fatigue and degraded weather conditions. Airspeed and altitude were identified as consistent, measurable and meaningful performance outcomes across a range of simulated flight scenarios. We were also able to begin demonstrating the ability to collect usable physiological data within operational type settings and the ability to use physiological data to predict meaningful flight performance outcomes. While these are significant accomplishments that will contribute to the development of an operator state monitoring system, they are only a small piece of the overall puzzle. Ongoing work will be required to fully establish such a system.												
15. SUBJECT TERMS operator state monitoring, physiological activity, aviation												
16. SECURITY CLASSIFICATION OF: <table border="1" style="width: 100%; border-collapse: collapse; font-size: 0.8em;"> <tr> <td style="width: 33%;">a. REPORT</td> <td style="width: 33%;">b. ABSTRACT</td> <td style="width: 33%;">c. THIS PAGE</td> </tr> <tr> <td style="text-align: center;">UNCLAS</td> <td style="text-align: center;">UNCLAS</td> <td style="text-align: center;">UNCLAS</td> </tr> </table>			a. REPORT	b. ABSTRACT	c. THIS PAGE	UNCLAS	UNCLAS	UNCLAS	17. LIMITATION OF ABSTRACT SAR		18. NUMBER OF PAGES 14	
a. REPORT	b. ABSTRACT	c. THIS PAGE										
UNCLAS	UNCLAS	UNCLAS										
			19a. NAME OF RESPONSIBLE PERSON Amanda Hayes, MS									
			19b. TELEPHONE NUMBER (Include area code) 334-255-6067									

This page is intentionally blank.

Summary

The project covered within this report addressed a number of areas related to the development of an operator state monitoring system for U.S. Army aviation. The table presented below provides a high-level overview of the activities that have been completed to date. While significant strides toward such a system were made through this project, many of them were exploratory or foundational in nature. Ongoing work will be required to achieve this system.

Objectives for the Present Project	Activities Completed Throughout the Project	Outcomes Achieved Throughout the Duration of the Project
	Survey of current Army aviators.	Confirmation of the perspective of fatigue and weather as prevalent factors impacting operational performance.
Development of an aviation composite performance index (CPI) to identify when performance becomes degraded, independent of maneuver type.	Evaluation of past U.S. Army Aeromedical Research Laboratory (USAARL) studies within the flight simulator.	Support for airspeed and altitude as meaningful performance outcome metrics.
	Retrospective analysis using data from previous USAARL studies.	Improved processes for data collection synchronization after determining inability to combine past datasets.
	Exploratory analysis using original study data and replication study data.	Identification of altitude predicting crash likelihood in simulated flights.
Refining algorithms and selecting variables to predict future changes in the CPI.	Evaluation of brain activity and heart activity data during fixed-wing live flight.	Demonstrated the ability to record usable physiological data during extreme maneuvering during live flight.
	Evaluation of brain activity monitoring devices for operational use.	Methods developed to measure brain activity in operational flight.
	Initial demonstrations of using physiological data to flight-related performance outcomes.	Demonstrated a brain-based metric in live, fixed-wing flight as sensitive to cognitive engagement and the use of brain-based metrics to predict crash likelihood in flight simulation.

This page is intentionally blank.

Acknowledgments

This research was supported in part by an appointment to the Research Participation Program at the U.S. Army Aeromedical Research Laboratory (USAARL) administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Army Medical Research and Development Command.

The authors would also like to acknowledge the USAARL Flight Systems Branch for their assistance throughout the study activities described here. The research pilots and branch chief, Mr. Jason Gerstner, were instrumental in the successes of these efforts. Additionally, the authors would like to acknowledge the professionalism and dedication of the research staff in the Warfighter Performance Group.

This page is intentionally blank.

Table of Contents

	Page
Summary	iii
Acknowledgments.....	v
Introduction.....	1
Establishing a Composite Performance Index	2
Timeframe.....	2
Current Effort.....	3
Objective 1: Development of an aviation CPI to identify when performance becomes degraded independent of maneuver type.	3
Identification of the conditions in which degraded performance occurs.	3
Measurable indicators of degraded performance.	6
When degraded performance leads to total failure.	8
Summary of Objective 1 Outcomes.....	8
Objective 2: Refining algorithms and selecting variables to predict future changes in the CPI	9
Assessing sensors in operational and operational-like settings.	9
Implementation of synchronization capabilities.	10
Evaluation of physiological data to predict performance outcomes.	10
Objective 2 Outcome Summary.....	11
Discussion	11
Recommendations	11
References.....	13

List of Figures and Tables

Figure 1. Combat deployments and the influence of environmental factors on performance.	4
Figure 2. Non-combat deployments and influence of environmental factors on performance.	5
Figure 3. Combat training centers and influence of environmental factors on performance.	5
Figure 4. Factors influencing performance for all environments.	6
Table 1. Frequencies of Studies Reporting Performance Outcomes Measured by Maneuvers/Phase of Flight	7
Table 2. Maneuvers/Phases of Flight and Performance Outcomes Reported in a Single Study	7

This page is intentionally blank.

Introduction

The U.S. Army Aeromedical Research Laboratory (USAARL) has been working toward an operator state monitoring (OSM) solution to maintain the safety and optimize the performance of Army aviators. Within this context, performance can be defined as an aviator's ability to successfully navigate a flight from beginning to end and successful completion of a mission. The intent for OSM is to use the physiological data from the aviator to predict performance-related outcomes. In their simplest form within this context, these outcomes can be binary: crash or successful landing; mission success or mission failure. In their more complex form, they can become continuous: altitude changes, airspeed changes, inputs to the controls, proper responses to radio calls, and other dynamic flight variables. Ultimately, in some form, the continuous outcomes contribute to the binary outcomes. A drastic change in altitude and airspeed can result in descending too fast and a crash landing. Missed radio calls could result in a failed mission. However, it is because of these various factors that contribute to success or failure that it is difficult to determine the aspects of performance to measure for development of such a system. Once reliable prediction of performance outcomes is achieved, monitoring the operator can be used to inform leadership of a Service Member's status for mission scheduling decisions, or feed into an adaptive automation and/or alerting system within the aircraft. Realizing such outcomes will require a robust OSM system that is reliably tied to meaningful performance outcomes and sensors selected that are capable of integration within operational settings.

Ongoing efforts at USAARL have been aimed at addressing these aspects of achieving an OSM system. Such efforts have included extensive collection of physiological data across various settings (e.g., cognitive tasks, the Multi-Attribute Task Battery [MATB], flight simulation) and with a variety of task manipulations to induce varying degrees of workload. These efforts have aided in laying the foundation for determining which physiological measures hold promise for reliably detecting performance decrements. For example, our past work has established a consistent relationship between frontal electroencephalography (EEG) activity and flight-performance degradation (Feltman et al., 2021). Other studies conducted at the laboratory have highlighted the importance of an individualized approach. One such effort explored individualized approaches to modeling physiological data for OSM (Vogl, O'Brien, et al., 2025), while another accounted for individual differences in responding to demanding flight tasks (Feltman et al., 2024).

While we have demonstrated *relationships* between flight performance measures and physiological measures, we had yet to demonstrate the ability to *predict* changes in performance from physiological measures. Given that successful flight requires the aviator to maintain several parameters, such as altitude and airspeed, we posited that the development of a composite performance index (CPI) would be beneficial for prediction purposes. Ideally, the CPI would be a value that evolves throughout the duration of a flight, continuously changing based on the inputs at that point in flight. However, to do so, we first needed to establish what aspects of performance the physiological measures need to predict to make the determination of whether an aviator is at-risk for performance decrements. In addition, the amount of time in which a reliable prediction of performance degradation can be made is needed to determine what interventions may be possible to mitigate risk. As such, the goal of the present work was to achieve the following: 1) to develop a CPI, and 2) to refine algorithms and select variables to establish the best offline prediction of future performance using the CPI through iterative experimentation.

Establishing a Composite Performance Index

A key component of OSM is identifying when an individual's performance is going to begin to degrade. Identifying what someone's physiological response looks like in the moments leading up to degraded performance will be key to ensuring some form of mitigation can be introduced *before* the decrements begin taking place. In many studies this is straightforward – a missed response, a slowed response, an incorrect response – but in applied research, it becomes messier. Missing something, being slow, or inputting something incorrectly does *not* always lead to meaningful problems with performance. In flight, sometimes small errors *do* have catastrophic outcomes, but that is relatively rare. A missed radio call may delay mission completion but not result in mission failure. An incorrect input into the system could also delay a mission but not necessarily result in a crash or a failed mission. This makes it difficult to determine *what* aspects of performance need to be predicted from physiological responses.

Within Army aviation, there are well-established standards to follow when executing various flight maneuvers. However, these differ across platforms (i.e., Chinook vs. Black Hawk) and within platforms, based on mission profiles and specific maneuvers (i.e., hover). Ideally, an OSM system capable of predicting performance degradations would be able to do so *regardless* of the specific platform and/or maneuver. Such an outcome would likely be composed of multiple performance indices (e.g., current altitude and airspeed), thereby creating a singular CPI. Changes in physiological measures, such as brain and heart activity, would predict the likelihood of that index changing and could then be used to trigger appropriate interventions (such as alerts or automation). Studies to date have examined individual performance outcomes, often without demonstrating practical significance of the outcomes. Take, for example, USAARL's previous flight simulation study where workload was manipulated while aviators performed terrain flights (Feltman et al., 2019). In this study, participants were instructed to maintain 200 feet above ground level (AGL) for both the low and high workload conditions. Results of the study yielded statistically significant findings, with greater altitude deviations in the high workload compared to the low workload condition. For the high workload condition, mean deviations were 25.25 feet, while in the low condition they were 19.85 feet. In live flight, however, it is typical of aviators to aim toward maintaining within 100 +/- AGL. USAARL's statistically significant finding of 25 feet deviations are operationally insignificant given the leeway of 100 feet. Therefore, altitude deviations alone are likely not a meaningful performance outcome.

Timeframe

Not only is *what* is being predicted important, but *how far in advance* is as well. In order to make any sort of meaningful intervention, there needs to be time for corrective action to take place. When using OSM to inform leadership about scheduling, the developed system has a wider timeframe to work with – changes in personnel scheduling generally has more flexibility. During live flight, however, the timeframe to act is less forgiving. In-flight, sometimes an aviator has only mere seconds to react and make a corrective action. Therefore, identifying the timeframe in which it can be predicted when that activity is trending toward failure is crucial for determining the ways in which OSM can be further developed and used. It is important to be able to identify the physiological measures that can provide the lead time needed to make a corrective action and sustain performance.

Current Effort

The purpose of the present report is to summarize USAARL activities over the fiscal year (FY) 24 (01 October 2023) through FY26 (30 September 2026) timeframe that served the purpose of addressing the objectives of 1) developing a CPI, and 2) refining algorithms and selecting variables to predict future changes in the CPI. The activities described here can be considered as an exploratory effort focused on the metrics to be included in furthering the development of a CPI, and ultimately, an OSM system.

Objective 1: Development of an aviation CPI to identify when performance becomes degraded independent of maneuver type.

In order to develop a CPI that is independent of maneuver type, several factors need to be established. These include: 1) identification of the conditions in which degraded performance occur; 2) measurable indicators of degraded performance; and 3) when degraded performance leads to total failure.

Identification of the conditions in which degraded performance occurs.

An OSM system will need to be capable of predicting the likelihood of performance degradation across a range of scenarios. To identify the variety of scenarios that place an aviator at risk for degraded performance, we conducted a survey of Army aviators. This survey effort included two iterations. The first included asking for feedback across three operational environments: 1) combat deployments, 2) non-combat deployments, and 3) combat training centers. The first iteration was conducted at two timepoints (14-16 May 2025 and 15-17 April 2026 at the Army Aviation Mission Solutions Summits, additionally online responses were received outside of the summit up to July 2026). The second iteration included reducing the question items by requesting participants to still evaluate performance across these three operational environments, but instead of asking the same questions for each environment, participants were asked to indicate in their responses whether one of the three was more of a factor in what they were reporting related to performance. The second iteration occurred August 2026.

First iteration - Survey of Army aviators.

To better understand factors influencing U.S. Army aviators' performance *today*, a survey was conducted. The survey was done in-person and virtually, with the survey items, methods, and preliminary outcomes summarized within a technical report (Feltman, Wilkins, Duffy, et al., 2026). The final survey had a total of 51 respondents (39 males, 12 females). Respondents were questioned on factors impacting performance across three operational settings: combat deployments, non-combat deployments, and combat training centers. The majority of participants reported total flight hours of 801 to 1500 hours ($n = 19$, 37.25%), while the second most frequent was 1501+ ($n = 18$, 35.29%). The remaining two hours categories, 201 to 500 hours and 501 to 800 hours, each included seven participants (13.73%). Regarding airframes flown, the majority reported H60s ($n = 37$, 72.55%); AH64s and CH47s tied for next most frequently reported ($n = 6$, 11.76%, each), one participant reported OH58 (1.91%) and one did not provide a response.

Participants were queried on whether they had been to each of the operational settings. Of those who responded to whether they have experienced a combat deployment, 27 reported one or more deployments, with 18 responding with zero. For non-combat deployments, 31 reported having been on one or more, and 10 responded with zero. Finally, 23 reported attending a combat training center one time or more, and 14 responded with none. Participants who had attended each of these settings once or more responded to survey on items on how various environmental factors impacted their performance. These factors included: weather, task saturation, fatigue, degraded visual environments (DVE), communications, and boredom. Figures 1-3 below summarize the frequency of responses to each item within each operational setting.

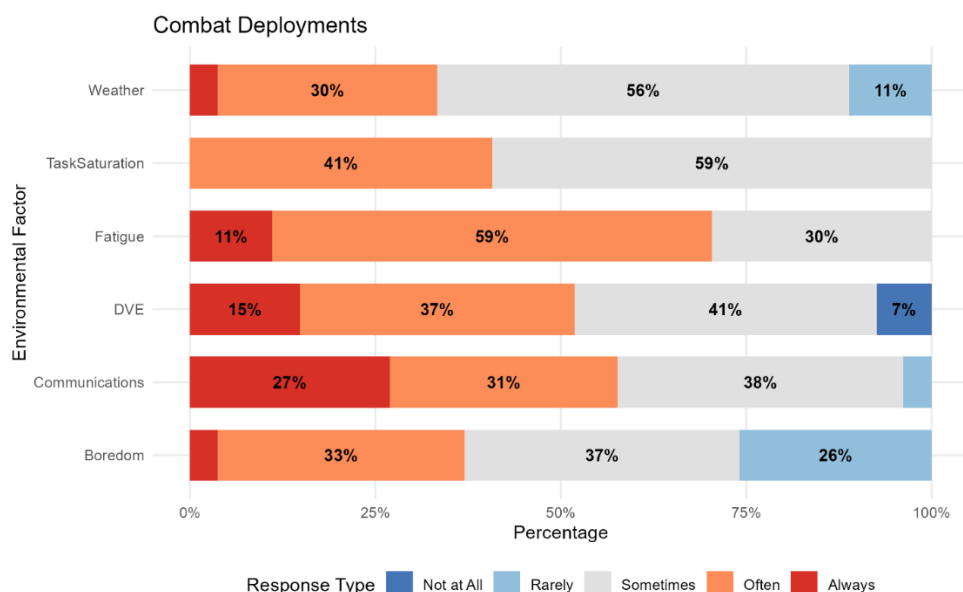


Figure 1. Combat deployments and the influence of environmental factors on performance.

This space is intentionally blank.

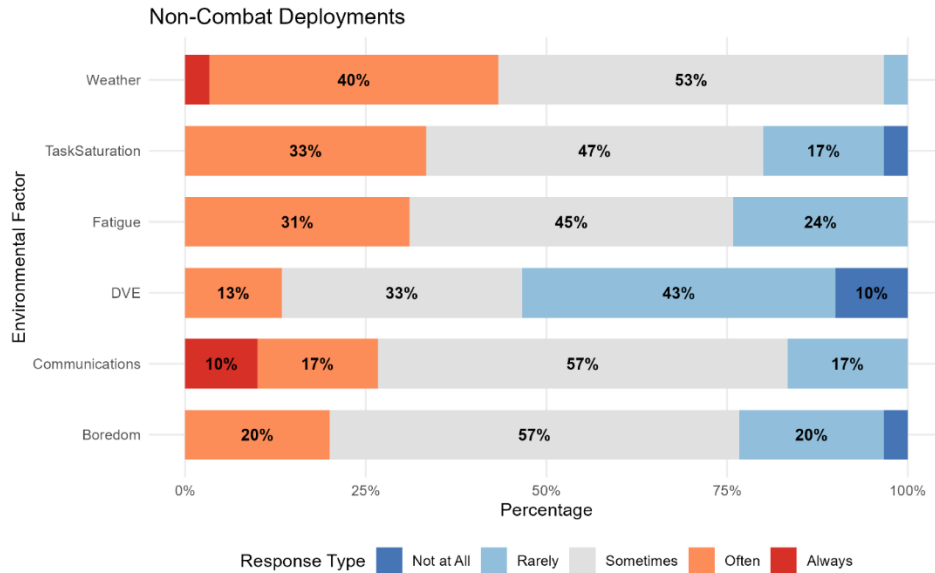


Figure 2. Non-combat deployments and influence of environmental factors on performance.

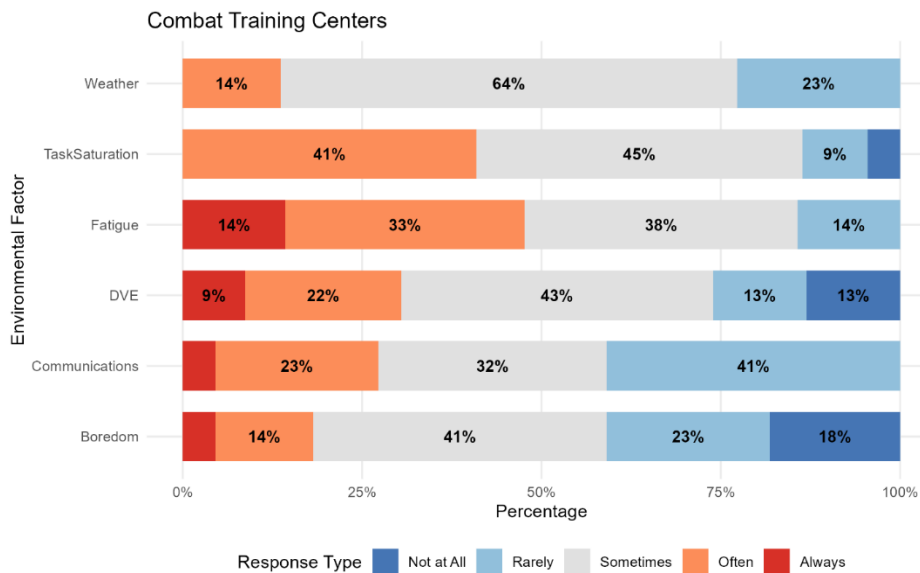


Figure 3. Combat training centers and influence of environmental factors on performance.

Overall, fatigue was most frequently rated as “Always” and “Often” within the combat deployment (combined ratings = 70%) and the combat training center (combined ratings = 47%) settings. For non-combat deployment settings, weather was the most prevalent factor (~40%).

Second iteration survey - Updated findings.

For the second iteration of the survey, a total of 18 participants responded. Of the 18, the majority were female ($n = 12$, 66.67%). The majority of participants reported flying the UH-60 ($n = 12$, 66.67%) and having flown in the mid-range of hours (201-500 hours, $n = 6$; 501-800 hours, $n = 7$). This condensed version of the survey asked participants to consider all three environments. From this sample, nine participants reported having gone on a combat deployment and non-combat deployment at least once, and 10 reported having gone to a combat training center at least once. In considering these environments, participants rated weather and fatigue as most frequently affecting performance. Weather received the greatest number of “Always” ratings. These two factors were highlighted in the previous survey iteration as well.

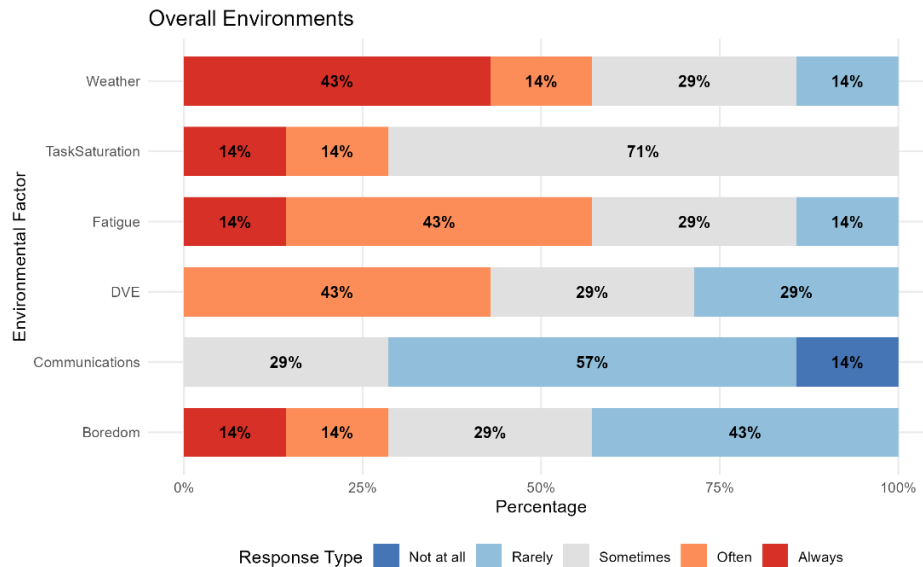


Figure 4. Factors influencing performance for all environments.

Measurable indicators of degraded performance.

To identify meaningful variables from simulated flight that indicate degraded performance, studies conducted at USAARL between FY15 and FY26 were examined. From this, two efforts were completed. The first was to summarize existing performance outcomes, and the second was to conduct a retrospective analysis of combined datasets.

Performance outcomes summary.

A total of 15 studies that used USAARL’s NUH-60 Black Hawk simulator were reviewed. The primary objectives of each study varied from neuromodulation performance enhancement to performance maintenance in DVE. Further, the majority of the studies included a workload manipulation as a part of the flight scenarios. Given the variability in the independent variables studied, study results cannot be used solely to estimate the value of any given flight maneuver and/or performance outcome. Thus, for each study, key information (flight maneuvers, performance outcomes, and study results were extracted, tabulated, and summarized (Tables 1 and 2).

Table 1. Frequencies of Studies Reporting Performance Outcomes Measured by Maneuvers/Phase of Flight

Maneuver/Phase of Flight	Performance Outcomes				
	Altitude	Heading	Rate of climb/descent	Airspeed	Lateral position
<i>Takeoff</i>	-	2	1	1	1
<i>Approach to hover</i>	2**	3*	-	2*	1*
<i>Hover</i>	5**	3**	-	-	2*
<i>En-route</i>	7***	2**	-	5***	1*
<i>Approach to landing</i>	1*	2*	1*	3**	2*
<i>Landing</i>	-	2**	-	2*	1

*Statistically significant in 1 study.

**Statistically significant in 2 studies.

***Statistically significant in 3 studies.

Table 2. Maneuvers/Phases of Flight and Performance Outcomes Reported in a Single Study

Maneuver/Phase of Flight	Performance Outcomes				
	Altitude	Heading	Rate of turn	Airspeed	Trim
<i>Pre-turn</i>	X	X	-	X	-
<i>Post-turn</i>	X	X	-	X	-
<i>Sidestep</i>	-	X*	-	X*	-
<i>Left turn</i>	X	-	X	X	X
<i>Right turn</i>	X*	-	X	X	X*

*Statistically significant

Additional performance outcomes that were included in a single study included glideslope, localizer instrument values, vertical speed, and tail velocity. Given their single use, it is challenging to draw any conclusions about the utility of these variables. They are all related in some capacity to the more commonly used outcomes, but their use may be advantageous if tailored to the needs of the research question.

The review of USAARL simulator flight performance research supports the utility of airspeed and altitude, especially when measured during en-route portions of flight, hovers, and approaches to landing.

Retrospective analysis.

Data from previous USAARL studies were reviewed to determine whether they could be combined for algorithm development purposes. The studies reviewed included a subset of seven studies from the previous 15 studies completed within the USAARL full-motion UH-60

simulator where flight performance data and physiological measures were collected. Upon reviewing the datasets, it was ultimately determined the data could not be reliably combined for the purpose of algorithm development. This was decided based on the following: data collection across multiple computers with separate clocks and timing formats, manually entered event markers that were susceptible to omission or delay, inconsistent recording timelines, proprietary device-derived computations, differing sampling rates, simulator upgrades, and variation in flight scenarios. These limitations prevented reliable post-hoc alignment of simulator, physiological, and event data across studies.

To address these limitations prospectively, the USAARL UH-60 simulator data collection architecture was modernized through integration of Lab Streaming Layer (LSL). LSL provides a common, network-time-synchronized framework for recording multimodal time-series data, including simulator performance data, physiological sensor streams, metadata, and, when applicable, in-the-loop subjective data. The resulting Extensible Data Format (XDF) files retain timing and metadata needed to correct offsets and network delays during analysis. This approach reduces reliance on manual synchronization procedures, creates auditable timing records, and establishes a more scalable foundation for pooling future studies to support OSM algorithm development. This effort has been described in a technical report by Vogl, Atchley, et al. (2025) and has since been implemented in all subsequent simulator studies. Thus, future efforts should be capable of pooling data to increase the size of datasets for algorithm development.

When degraded performance leads to total failure.

Although we were unable to combine multiple datasets to create a larger dataset for machine learning purposes, there were two studies we were able to use for exploring different statistical techniques. One of the two studies was designed to be a replication of the other. Therefore, although they suffered from similar limitations with synchronization of data streams, they at least contained enough similarities to be combined. Furthermore, these two datasets included instances of “crashes” during simulation. Thus, there was a built-in, clear-cut indicator of performance success vs. performance failure. These datasets were then used to develop classification models predicting crash outcome. The details of this process are outlined in a manuscript and summarized here.

One dataset was used to train the models and the other to validate. From this analysis, we were able to use the changes in altitude to effectively predict nearly half of the crashes in the dataset. While on the surface that appears to be unsubstantial, given the crashes were a relatively rare event, this accuracy rate is consequential. Moreover, it is anticipated to improve as we add in additional predictor variables with future model development. This paper is currently in preparation for journal submission (Persson & Feltman, 2026).

Summary of Objective 1 Outcomes

Taken together, the efforts undertaken to address Objective 1 have improved our knowledge of the conditions that increase the likelihood of performance degradation such as fatigue and weather, identified measurable indicators of degraded performance (airspeed and altitude during critical phases of flight), and narrowed the scope of variables to examine when looking into total performance failure. While we determined it was not feasible to develop a CPI

based on current available data, the outcomes from this effort will inform future study design and data evaluation.

Objective 2: Refining algorithms and selecting variables to predict future changes in the CPI

Once it was determined that the development of a CPI was not feasible at this time, the methods to address this objective were slightly modified. As we were unable to develop a CPI to then assess various predictive algorithms, we instead focused on the selection of variables. Given the intent of the OSM system is to use physiological data along with aircraft state data to predict future decrements, we focused on several factors related to determining physiological variables for use in OSM. These factors included: 1) assessing sensors in operational or operational-like settings, 2) implementation of synchronization capabilities within the research simulator, and 3) evaluation of physiological data in predicting performance outcomes.

Assessing sensors in operational and operational-like settings.

Moving wearables from the laboratory to the field requires sensors that can reliably collect data of a usable quality in field settings. In rotary-wing aviation, this means they need to be able to be worn with the aviator's safety equipment and be capable of withstanding exposure to varying temperatures and vibrations. Previous research has identified EEG and electrocardiogram (ECG) as promising measures to continue exploring (Warfighter Performance Group, 2024). To this end, the following efforts were completed: evaluation of EEG and ECG data in fixed-wing live flight, integration of an EEG system beneath the aviator's helmet during simulated flight, and the evaluation of a small form EEG device. These efforts are briefly described here as each has its own corresponding technical report and/or published manuscript.

Evaluation of EEG and ECG data during fixed-wing live flight.

EEG and ECG were collected during an upset prevention and recovery training event (Yue et al., 2025). From this event, the research team was able to demonstrate the feasibility of collecting this data in such a setting. This included developing a process of recording during actual flight and subsequently processing the data for analysis. The outcomes from this effort demonstrated the ability to record usable EEG and ECG physiological data during extreme maneuvering during live flight.

Physiological monitoring beyond the laboratory.

While we were successful collecting usable EEG data during the fixed-wing live flight, we still do not have a device that is easy to implement and integrates with aviators' safety equipment. Two efforts have been completed to address this: 1) the evaluation of a small form EEG system (Atchley et al., 2026), and 2) the integration of EEG into the aviator's helmet during simulated flight (Feltman, Wilkins, Andres, et al., 2026).

From these evaluations we have been able to identify recommendations for requirements for similar systems to be capable of use in operational flight settings. Further, these findings are now being used to inform companies who are working to develop research solutions to more comfortably fit a device within the helmet to measure brain activity. In addition, we have taken

the lessons from this evaluation and have made modifications to improve comfortability. This will allow us to collect data from this research system in-flight for upcoming studies.

Implementation of synchronization capabilities.

Another critical piece of the work being done is to the necessity to collect the physiological data in synchronization with flight data. The initial integration of LSL into the simulator has been a critical step in achieving this. Ongoing efforts are working to replicate the simulator integration into the USAARL research helicopter for further OSM-related efforts.

Evaluation of physiological data to predict performance outcomes.

A preliminary live flight evaluation demonstrated the feasibility of acquiring usable EEG data during upset prevention and recovery training (Yue, Duffy, et al., 2026). Two pilots completed commercial training events while EEG was recorded. An EEG-derived engagement index showed partial evidence of higher engagement during maneuver execution than during instruction or observation. In addition, a cumulative engagement index (CEI) was developed to characterize longer-duration changes in engagement during repeated stall-recovery maneuvers. Across the final maneuver repetitions, CEI patterns were consistent with decreasing cognitive demand as pilots gained familiarity with the task.

These results are preliminary and observational. The small sample, real-world training conditions, and imperfect event synchronization preclude attribution of the observed EEG changes to specific cognitive processes. Nonetheless, the findings provide evidence that EEG-derived measures can capture task-related cognitive dynamics during demanding live flight training and support continued development of synchronized in-flight physiological-data collection methods.

A separate proof-of-concept analysis evaluated whether global EEG engagement indices and flight-context variables could provide time-resolved estimates of simulated-flight crash risk (Yue, Feltman, et al., 2026). Twelve aviators completed a challenging, high-fidelity Black Hawk simulator route; five flights ended in terrain collisions. To minimize confounding from differences in routes and flight durations, crash and non-crash trajectories were aligned using flight-context features, including altitude, longitudinal position, cumulative travel distance, and elapsed time. A penalized time-varying Cox model incorporating three global EEG engagement indices and terrain elevation showed better fit than a null model and demonstrated strong apparent internal discrimination; however, the modest sample limits confidence in generalizability.

Within this model, the alpha-normalized engagement index, defined as $1/\alpha$ power, was negatively associated with instantaneous crash hazard. This pattern is consistent with the interpretation that lower alpha power reflects greater externally oriented attention and visuospatial engagement. Kalman-smoothed hazard trajectories for crash and non-crash flights separated approximately 56 seconds before impact. This value should not be interpreted as an operational alarm threshold; rather, it provides preliminary group-level evidence that physiological and contextual indicators may diverge sufficiently before a performance failure to warrant further investigation of potential intervention windows.

Objective 2 Outcome Summary

Although a maneuver-independent CPI could not be derived from the currently available data, the project advanced three components of OSM development. First, live flight work demonstrated that EEG-based measures can be collected and evaluated during demanding training events. Second, implementation of an LSL-based synchronization architecture addressed critical legacy barriers to integrating physiological and simulator data across future studies. Third, simulated-flight modeling demonstrated that interpretable EEG engagement metrics, when combined with operational context, may support time-resolved estimation of risk preceding a defined performance failure. Collectively, these efforts shift the program toward prospective, synchronized, multimodal datasets capable of supporting more rigorous development and validation of OSM algorithms.

Discussion

While the objectives were not fully achieved in the current effort, substantial steps toward addressing several of the key elements of developing an OSM system were achieved. Critically, the demonstrations of synchronized data collections and capturing meaningful physiological indices during live flight are crucial achievements. By establishing synchronization capabilities, we can better examine our physiological data alongside what is actually happening during flight. As we move this capability into the research helicopter, we will be able to build on our achievement of live flight physiological data collection. Furthermore, we were able to confirm amongst current aviators what factors remain at the forefront of putting them at risk for degradation. This confirmation coupled with our literature review and retrospective analyses provide the foundation needed to continue development in this area.

Recommendations

From the current effort, we can describe some next steps that will be needed for the development of an OSM system to continue. These are in no particular order, as much of the work will need to be done concurrently to achieve the end-state of a functioning OSM system.

Physiological data needs to be collected during live flight operational scenarios with aviators' safety equipment worn. One key next step is to get the technology out of the simulation environment and into live flight. This is a necessity to work out the algorithms that will be needed to process physiological signals collected in such an environment. While we have been able to demonstrate that usable data can be collected beneath the helmet in simulation, and, separately, usable data during extreme maneuvering in fixed-wing flight, further demonstration in U.S. Army operational flights is needed. The development of proper data processing techniques will require more datasets across a variety of environments.

Synchronization of helicopter flight performance data with flight worthy physiological devices is needed. While collecting physiological data in-flight for the purposes of developing processing techniques is critical, so is the flight performance data. To get an OSM system to a usable state for predicting performance degradations, we will need to know *what* the aviator is doing when their physiological signals are changing. This will require the replication of our synchronization capabilities within the simulator into our research helicopter.

Determination should be made of whether signals captured in simulation are representative of those in live flight. Although the goal is to move more of the OSM development into live flight, there will be scenarios where simulation is still required. Such scenarios include when resources are limited (live flight is more expensive than simulation), tasks need to be isolated (more difficult to control in live flight), or the tasks of interest are too dangerous to perform in live flight. However, we do not have concrete evidence of whether what is captured in simulated flight is representative of that in live flight. Direct comparisons should aid in resolving this item.

Collection of physiological data across a wide range of settings, with different individuals is needed. More data is most certainly needed to train any models to be used in performance prediction. Part of this will require collecting data across a range of settings, not just the “pure” experimental ones. Getting devices into field settings will go a long way in aiding in the understanding of what these signals look like in this population across time and place.

Identifying which physiological metrics and at what level of granularity will be required is important. Further work is required to determine whether a wholly individualized approach to OSM is needed, or if group level data can be used to build out and inform models. Historically, we have used both approaches and ongoing work will continue to explore both.

References

- Atchley, A. J., Duffy, M., Feltman, K., Basso, J., & Yue, X. (2026). *Evaluation of signal quality and comfort for wearable EEG device in the UH-60 flight simulator* (Report No. USAARL-TECH-TR--2026-31). U.S. Army Aeromedical Research Laboratory.
- Feltman, K., Kelley, A., Bernhardt, K., Britt, T., & Mathews, C. (2019). *Real-time operator state monitoring during periods of high and low workload across four workload domains* (Report No. 2019-12). U.S. Army Aeromedical Research Laboratory.
- Feltman, K., Wilkins, J., Andres, K., Basso, J., & Yue, X. (2026). *Evaluation of a commercial EEG system beneath the aviator's helmet: Data quality and comfortability* (Report No. USAARL-TECH-FR--2026-06). U.S. Army Aeromedical Research Laboratory.
- Feltman, K., Wilkins, J., Duffy, M., Wolf, S., Wood, M., & Gerstner, J. (2026). *Identifying the top factors and stressors impacting flight performance: A Survey of Army aviators*. (Report No. USAARL-TECH-TR--2026-03). U.S. Army Aeromedical Research Laboratory.
- Feltman, K. A., Kelley, A. M., Bernhardt, K., Basso, J., & Morabito, C. (2021). *Psychophysiological indicators of aviator flight performance for operator state monitoring* (Report No. USAARL-TECH-FR--2021-17). U.S. Army Aeromedical Research Laboratory.
- Feltman, K. A., Vogl, J. F., McAtee, A., & Kelley, A. M. (2024). Measuring aviator workload using EEG: An individualized approach to workload manipulation. *Frontiers in Neuroergonomics*, 5. <https://doi.org/10.3389/fnrgo.2024.1397586>
- Persson, I., & Feltman, K. A. (2026). *Predictive analysis of simulator crashes using machine learning methods and flight performance data* [Manuscript under review]. U.S. Army Aeromedical Research Laboratory.
- Vogl, J., O'Brien, K., & St. Onge, P. (2025). One size does not fit all: a support vector machine exploration of multiclass cognitive state classifications using physiological measures. *Frontiers in Neuroergonomics*, 6. <https://doi.org/10.3389/fnrgo.2025.1566431>
- Vogl, J. F., Atchley, A., Persson, I., Mackie, R., & Feltman, K. A. (2025). *Towards data synchronization and real-time processing methodologies for operator state monitoring research at USAARL* (Report No. USAARL-TECH-TR--2025-27). U.S. Army Aeromedical Research Laboratory.
- Warfighter Performance Group. (2024). *Choosing the right gear: A guide to physiological devices for monitoring operator states of aviators* (Report No. USAARL-TECH-HB--2024-34). U.S. Army Aeromedical Research Laboratory.

- Yue, X., Andres, K., Duffy, M. J., Rodriguez, R., Aura, C. J., & Feltman, K. A. (2025). Measuring pilot physiology during in-flight training and implications for real-time monitoring. *Aerospace Medicine and Human Performance*, 96(5), 447–452. <https://doi.org/https://doi.org/10.3357/AMHP.6596.2025>
- Yue, X., Duffy, M., & Feltman, K. (2026). A method to quantify cognitive state in flight through a brain-based cumulative engagement index [Manuscript in press]. *Aerospace Medicine and Human Performance*.
- Yue, X., Feltman, K. A., Vogl, J. F., & Kelley, A. (2026). Time-varying EEG survival modeling predicts aviation crash risk in simulated flight [Manuscript under review]. *Scientific Reports*.

U.S. ARMY AEROMEDICAL RESEARCH LABORATORY



FORT RUCKER, ALABAMA

Optimizing

**HUMAN PROTECTION
AND PERFORMANCE**
since 1962



U.S. ARMY



FCC



MRDC