LIMITATIONS OF FATIGUE RISK PERFORMANCE MODELING

Sarah A. Chabal, Ph.D.[1,2], Emily M. Moslener[1,2], Rachel R. Markwald, Ph.D.[1]

[1]Warfighter Performance Department, Naval Health Research Center, San Diego, CA, USA; [2]Behavioral Health & Readiness Division, Health & Services Sector, Leidos Inc., San Diego, CA, USA

NAVY MEDICINE RESEARCH & DEVELOPMENT

BACKGROUND

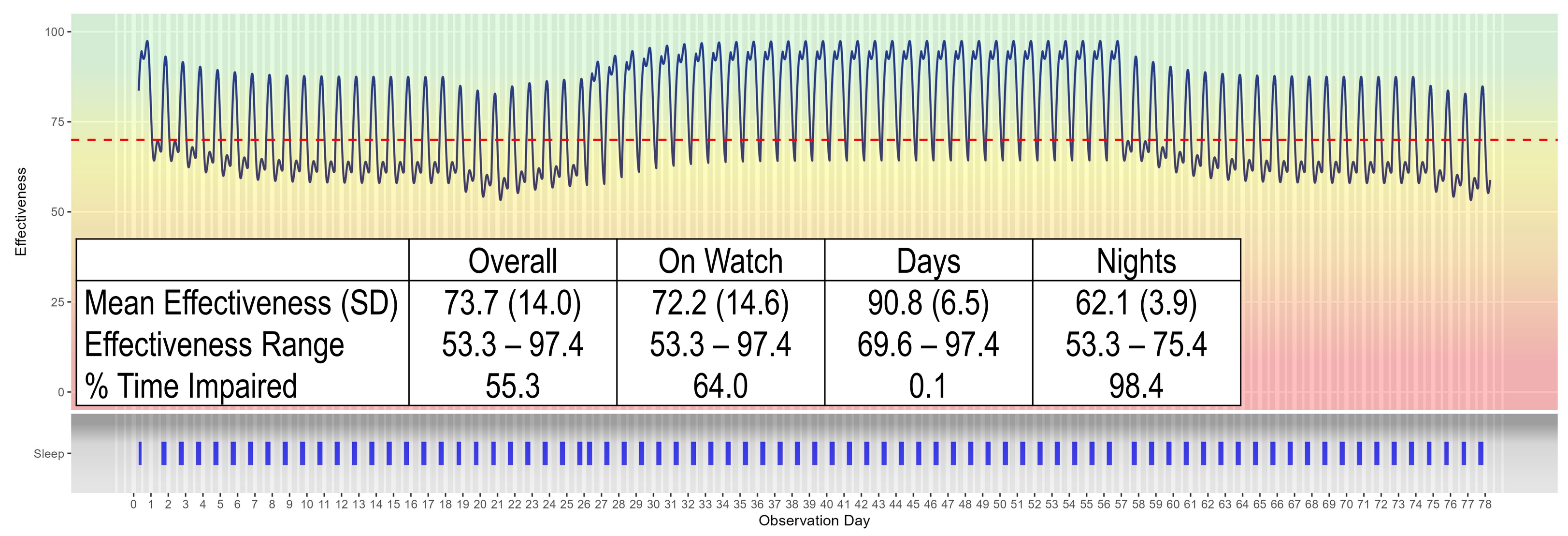
- Assessment of military operators' cognitive performance is crucial to ensure safe, successful operations. Unfortunately, real-time performance assessments interrupt military task performance.
- The Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) model[1] uses sleep history/ circadian cycles to predict cognitive effectiveness. Predictions are validated against neurobehavioral performance[2] and correlate with the prevalence of safety accidents.[3]
- SAFTE data are used by military leadership in contexts including safety assessments of training exercises, evaluations of countermeasure effectiveness,[4] watchbill management,[5] and mishap evaluations.[6]
- SAFTE models can only be as accurate as the data used to build them; however, there is currently no guidance on the level of data granularity that should be used for SAFTE model construction.

METHODS

- Data were collected from a shore-based U.S. Navy command that followed a four-section, 12-hour watchbill with a Panama schedule and 28-day rotation
- Sleep schedules were constructed from three sources:
  - Idealized sleep based on work schedules alone (assumptions: 7 hrs sleep/ day)
  - Idealized sleep based on group-level survey data (assumptions: 7.6 hrs sleep/ day on day shift and days off, 6 hrs sleep/ day on night shift)
  - Objective sleep derived from 2 individuals who each provided 82 days of actigraphy data (Oura Ring)
- Four performance models were constructed using open-source SAFTE in R (SAFTEr) and compared using multi-level regression models

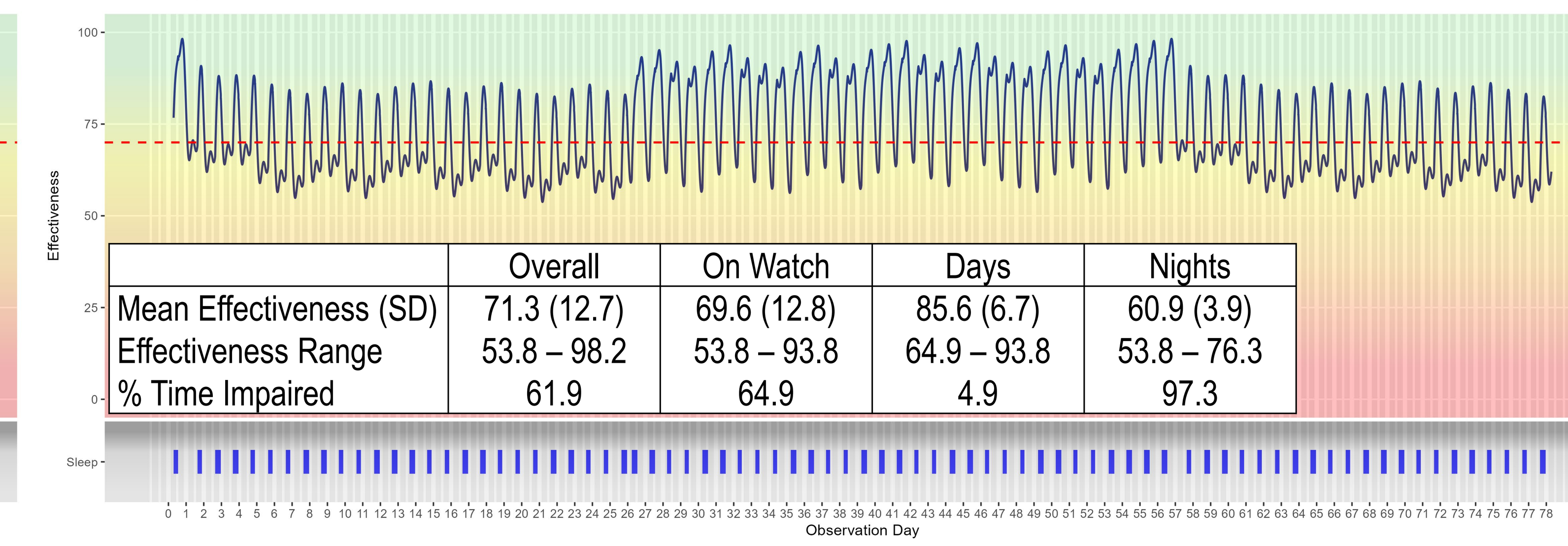
RESULTS

Schedule-Based SAFTEr Model



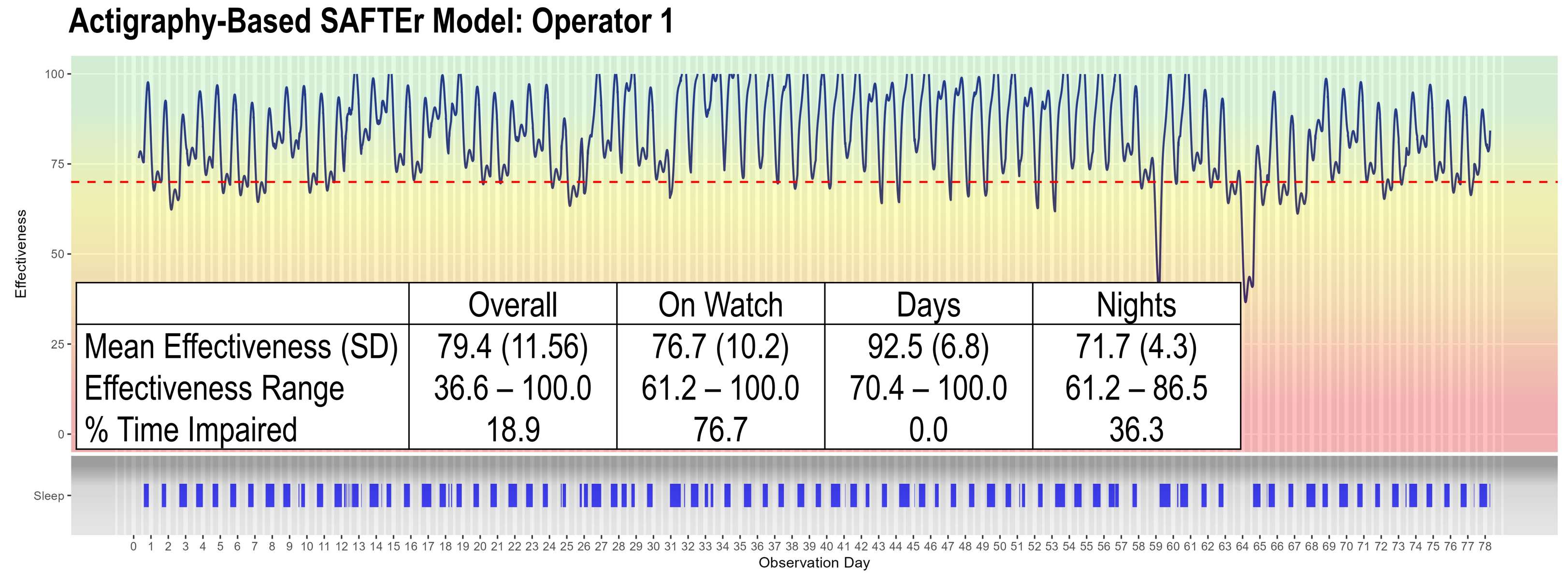
	Overall	On Watch	Days	Nights
Mean Effectiveness (SD)	73.7 (14.0)	72.2 (14.6)	90.8 (6.5)	62.1 (3.9)
Effectiveness Range	53.3 – 97.4	53.3 – 97.4	69.6 – 97.4	53.3 – 75.4
% Time Impaired	55.3	64.0	0.1	98.4

Survey-Based SAFTEr Model



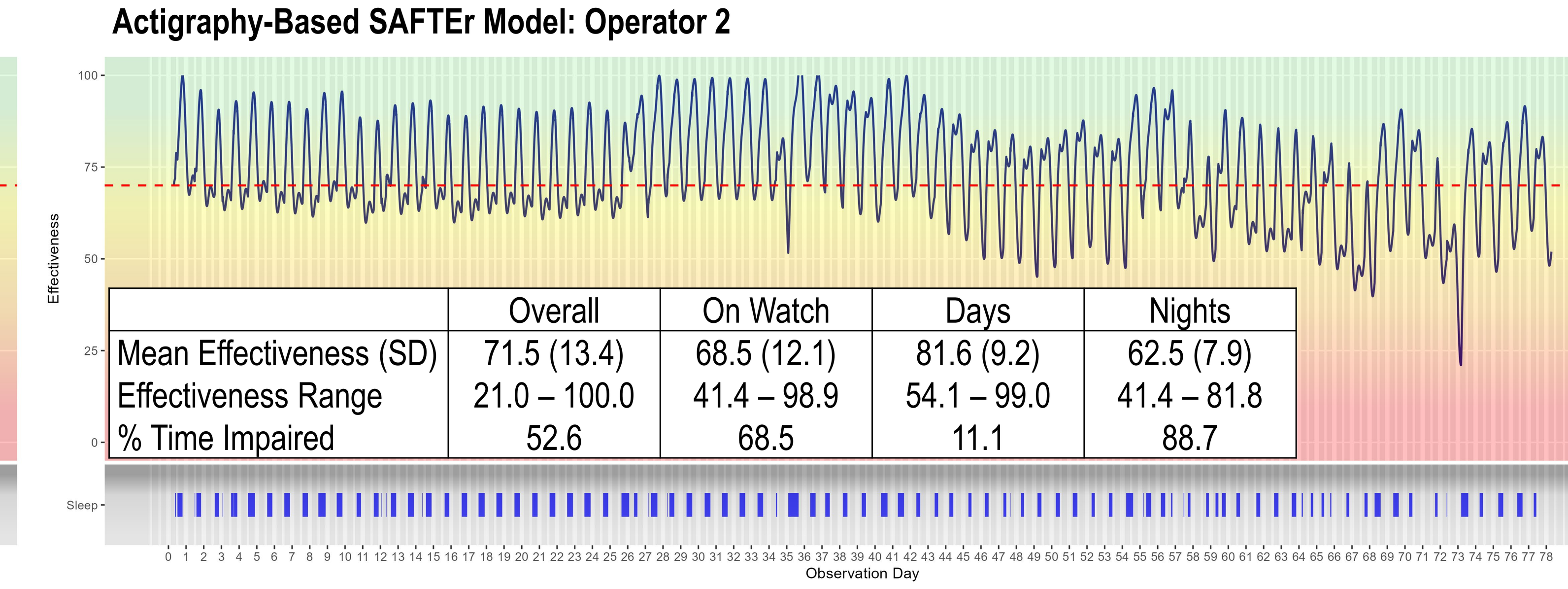
	Overall	On Watch	Days	Nights
Mean Effectiveness (SD)	71.3 (12.7)	69.6 (12.8)	85.6 (6.7)	60.9 (3.9)
Effectiveness Range	53.8 – 98.2	53.8 – 93.8	64.9 – 93.8	53.8 – 76.3
% Time Impaired	61.9	64.9	4.9	97.3

Actigraphy-Based SAFTEr Model: Operator 1



	Overall	On Watch	Days	Nights
Mean Effectiveness (SD)	79.4 (11.56)	76.7 (10.2)	92.5 (6.8)	71.7 (4.3)
Effectiveness Range	36.6 – 100.0	61.2 – 100.0	70.4 – 100.0	61.2 – 86.5
% Time Impaired	18.9	76.7	0.0	36.3

Actigraphy-Based SAFTEr Model: Operator 2



	Overall	On Watch	Days	Nights
Mean Effectiveness (SD)	71.5 (13.4)	68.5 (12.1)	81.6 (9.2)	62.5 (7.9)
Effectiveness Range	21.0 – 100.0	41.4 – 98.9	54.1 – 99.0	41.4 – 81.8
% Time Impaired	52.6	68.5	11.1	88.7

Note. Overall Effectiveness refers to projected performance effectiveness during all waking times, regardless of whether the Operator is at work. Effectiveness on Watch refers to projected performance effectiveness while at work collapsed across all watch sections. Effectiveness on Days refers to projected performance effectiveness while standing the days watch. Effectiveness on Nights refers to projected performance effectiveness while standing the nights watch. Impairment is defined as effectiveness below 70% (correlated with behavioral performance at a blood alcohol content of 0.08) and is represented by a red dashed line on all figures; % Time Impaired reflects the percentage of each period spent with impaired performance.

- All models produce statistically-different outcomes (all *ps* < 0.01)
- On-watch performance predictions vary by 5.2% between work schedule vs. survey data, and vary by 8.2% between individuals
- Survey-based models underestimate mean on-watch performance for Operator 1 by 7.2% and overestimate mean on-watch performance for Operator 2 by 1.1%
- Actigraphy data demonstrate that sleep occurs during unexpected times (Operator 1: 0.1% of sleep is during scheduled work times; Operator 2: 2.1% of sleep is during scheduled work times)
- Mean performance is similar across models; actigraphy-based models capture a larger range of performance

REFERENCES

1. Hursh, S.R., et al., Fatigue models for applied research in warfighting. Aviation, Space, and Environmental Medicine, 2004. 75(3): p. A44-A53. 2. Roma, P.G., et al., Flight attendant work/rest patterns, alertness, and performance assessment: Field validation of biomathematical fatigue modeling. 2012, Federal Aviation Administration: Washington, DC. 3. Hursh, S.R., J.F. Fanzone, and T.G. Raslear, Analysis of the relationship between operator effectiveness measures and economic impacts of rail accidents. 2011, U.S. Department of Transportation: Washington, DC. 4. Chabal, S., et al., Personal light treatment devices: A countermeasure to improve sleep, fatigue, and circadian misalignment in an operational setting. SLEEP Advances, 2025. 6. 5. Chabal, S., R.R. Markwald, and E.D. Chinoy, Life onboard a submarine: Sleep, fatigue, and lifestyle behaviors of Sailors on a circadian-aligned watchstanding schedule. Applied Ergonomics, 2024. 119(104321). 6. Miller, N.L., P. Matsangas, and L.G. Shattuck, Fatigue and its effect on performance in military environments, in Performance Under Stress, P.A. Hancock and J.L. Szalma, Editors. 2008, Ashgate Publishing: Burlington, VT.

CONCLUSIONS

- SAFTE model outputs are vulnerable to individual differences in sleep habits despite shared work schedules.
- Schedule- and survey-based models represent performance under steady, ideal circumstances.
- Actigraphy-based models capture extremes and large performance changes due to acute sleep loss, and capture sleep that occurs during unexpected times.
- Models based on group-level sleep estimates can inform schedule management decisions (e.g., timing/ frequency of operations).
- Individualized, actigraphy-based models should be used to fine-tune work schedules and inform readiness decisions.

Disclaimer: Some authors are military service members or employees of the U.S. Government. This work was prepared as part of their official duties. Title 17, U.S.C. §105 provides that copyright protection under this title is not available for any work of the U.S. Government. Title 17, U.S.C. §101 defines a U.S. Government work as work prepared by a military service member or employee of the U.S. Government as part of that person's official duties. This work was supported Naval Medical Forces Pacific and Naval Advanced Medical Development under work unit no. 62513. The views expressed in this work are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government. The study protocol was approved by the Naval Health Research Center Institutional Review Board in compliance with all applicable federal regulations governing the protection of human subjects. Research data were derived from approved Naval Health Research Center Institutional Review Board protocol number NHRC.2025.0007.