INSTITUTE FOR DEFENSE ANALYSES

Sequential Space-Filling Designs for Modeling & Simulation Analyses

John T. Haman, Project Leader

Anna R. Flowers

OED

November 2024

IDA Product ID 3003752

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22305

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

About This Publication
This work was conducted by the Institute for Defense Analyses (IDA) under the central research program, C9082, "Statistics and Data Science Working Group". The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments
The IDA Technical Review Committee was chaired by Dr. Heather Wojton and consisted of Dr. Jason Sheldon, Dr. Miriam Armstrong, Dr. Brian Vickers, Dr. Rebecca Medlin, Dr. Dhruv Patel, Dr. Curtis Miller, and Dr. Keyla Pagan-Rivera, from the Operational Evaluation Division.

For more information:
Dr. John Haman, Project Leader
jhaman@ida.org • (703) 845-2132

Dr. Heather Wojton, Director, Operational Evaluation Division
hwojton@ida.org • (703) 845-6811

Rigorous Analysis | Trusted Expertise | Service to the Nation

INSTITUTE FOR DEFENSE ANALYSES

IDA Product ID 3003752

Sequential Space-Filling Designs for Modeling & Simulation Analyses

John T. Haman, Project Leader

Anna R. Flowers

This page intentionally left blank

Executive Summary

This presentation provides an overview of sequential space-filling designs, with the focus on designs that are most suitable for the test and evaluation community.

There are few recommended methods to help testers plan efficient modeling and simulation studies. Space-filling designs are a rigorous choice, but a drawback of standard methods is that they require the final sample size to be selected prior to testing. More efficient testing can be completed by using sequential design strategies, which choose test points without knowledge of the final sample size. Using sequential designs can prevent oversampling and help to augment poorly designed tests.

Space-filling designs are a type of experimental design that are used for modeling and simulation (M&S) studies. Unlike traditional experimental design methods, they provide good coverage of the operational space, and are thus typically better at capturing model trends. These designs are selected based on some space-filling criterion, of which there are many. Sequential space-filling designs, the topic of interest in this presentation, are sequential designs which are space-filling at every step of the testing process.

In addition to sequential space-filling designs, there is another type of sequential design that is potentially useful to the test and evaluation community. This is called adaptive sampling (or active learning in the machine learning community), and it selects new test points based on model output. Rather than relying on coverage of the operational space, adaptive sampling chooses test points where the model is least certain; either where there is no data, or where there are complicated response dynamics.

Categorical variables are common in test and evaluation data. Although many statistical methods exist to sequentially design experiments, most cannot accommodate categorical variables. Additionally, many data sets contain binary outputs, such as hit/miss. This is another important distinction because it affects the selection criterion used for adaptive sampling. One key aspect of this presentation is the identification of sequential design techniques that allow both categorical variables and binary outcomes.

IDA researched sequential space-filling designs and adaptive sampling techniques and compiled the three methods most likely to be of use for the test and evaluation

community. We then performed a simulation experiment to compare the performance of those three methods.

In our simulation, we studied a Bernoulli outcome scenario with categorical factors, and applied three sequential space-filling design methods: adaptive sampling, sequential MaxPro[1], and sequential sliced Latin hypercube designs. Simulation results indicate that all three methods performed similarly. Based on this finding, and due to its flexibility, we recommend that testers use the MaxPro method for sequential space-filling designs in M&S studies.

---

[1] Joseph, V. Roshan, Evren Gul, and Shan Ba. "Designing Computer Experiments with Multiple Types of Factors: The MaxPro Approach." *Journal of Quality Technology* 52, no. 4 (October 1, 2020). https://doi.org/10.1080/00224065.2019.1611351.

Sequential Space-Filling Designs for Modeling & Simulation Analyses

Anna Flowers

Mentors:

John Haman, Rebecca Medlin, Curtis Miller, Keyla Pagán-Rivera, and Dhruv Patel, OED

Justin Krometis, Virginia Tech National Security Institute

August 1, 2024

Institute for Defense Analyses

730 East Glebe Road ● Alexandria, Virginia 22305

Bottom Line Up Front

Motivation: Few recommended methods exist to help testers plan efficient M&S studies

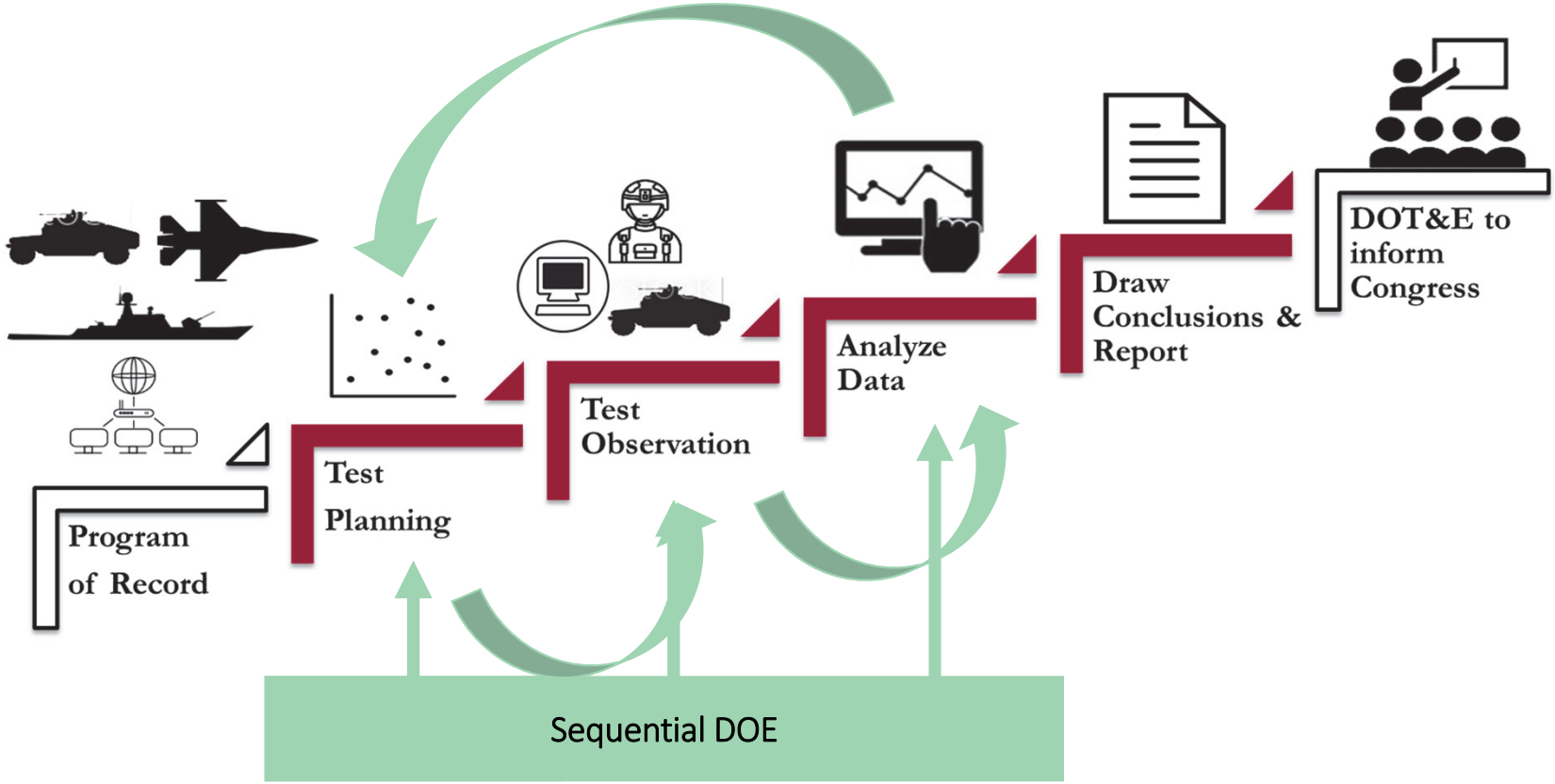
- SFDs are a rigorous choice, but standard methods require sample size to be selected prior to testing
- Efficient testing can be achieved by using sequential design: choosing test points without knowledge of the final sample size
- Sequential designs prevent oversampling and can be used to augment poorly designed tests

Project Goal: Review existing sequential SFD methods and provide recommendations on the best designs for scenarios we expect to see in T&E

Outcome: We used a simulation study to compare sequential SFDs and will publish a literature review of them with recommendations for the T&E community

SFD: Space-filling design; M&S: Modeling and Simulation; T&E: Test and Evaluation

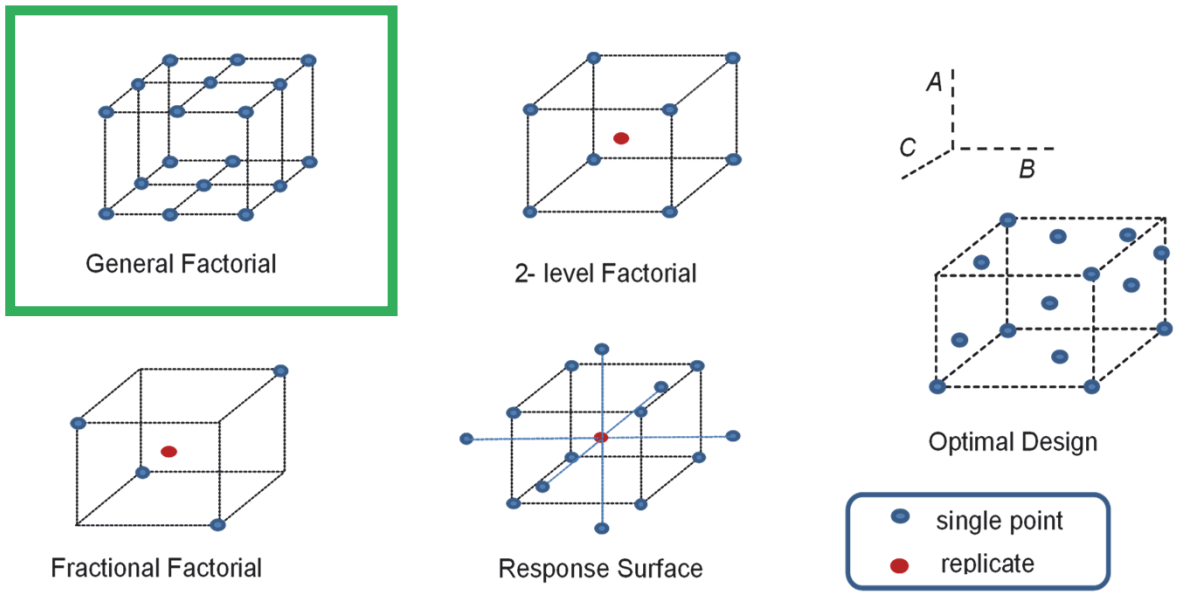
IDA's Operational Evaluation Division (OED) provides technical and analytical support to DOT&E



DOT&E: Director, Operational Test and Evaluation; DOE: Design of Experiments

Design of experiments is used to choose the “best” test points

- DOE provides a numerical measure of design quality through metrics such as power
- Classical DOE methods favor replicates and points near the boundaries

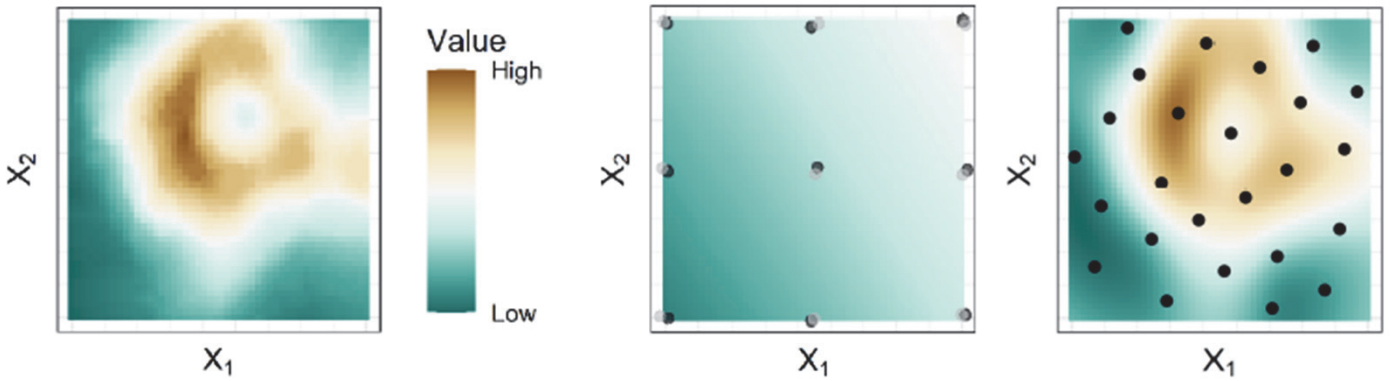


DOE: Design of Experiments

Image from Haman MS Validation IDA Analyst Course

Space-filling designs are often used for M&S

- SFDs provide good coverage of the operational space
- They are typically better at finding model trends than classical design strategies



Ground Truth

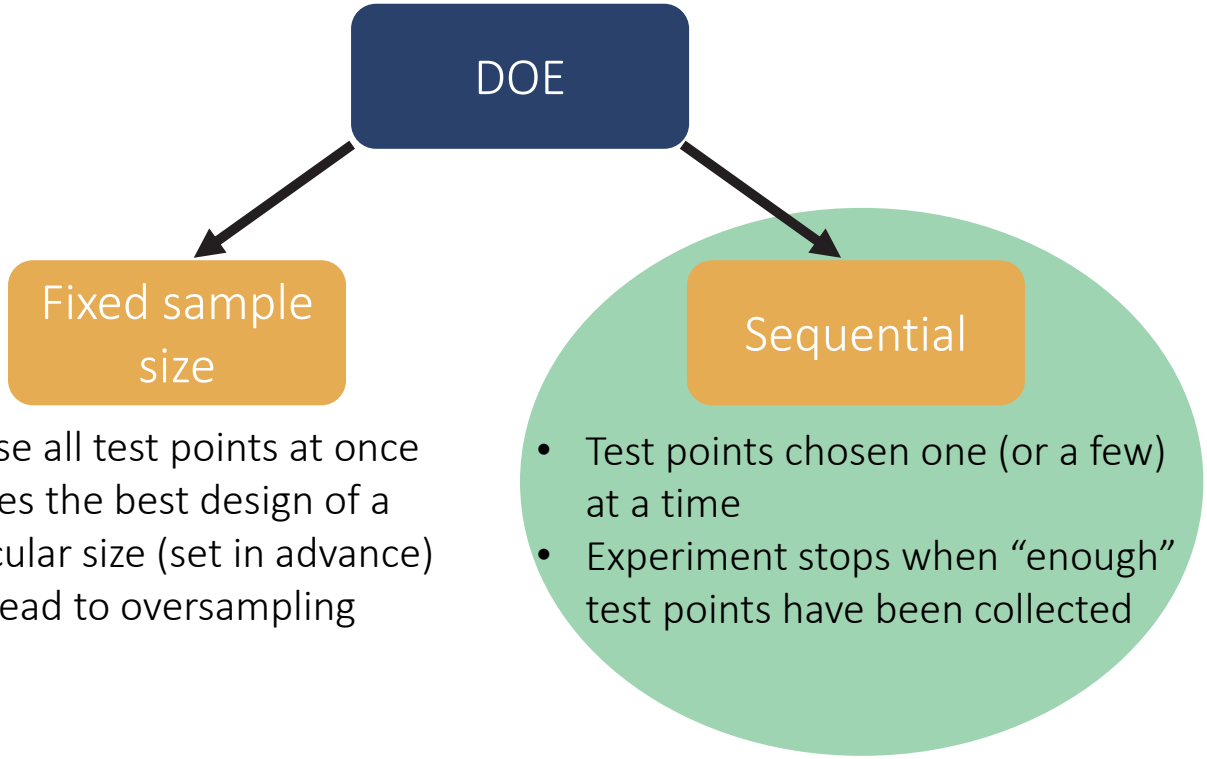
Classical design (n=25)

SFD (n=25)

Figure from Wojton et. al (2021)

M&S: Modeling and simulation; SFD: Space-Filling Design

Test points can be collected all at once or sequentially

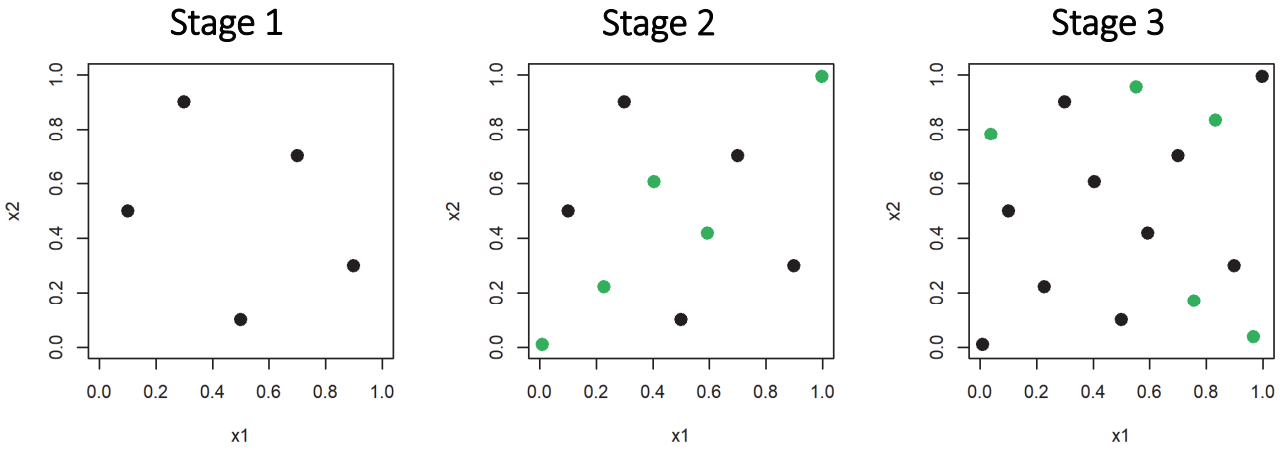


Sequential designs are attractive because of their flexibility.

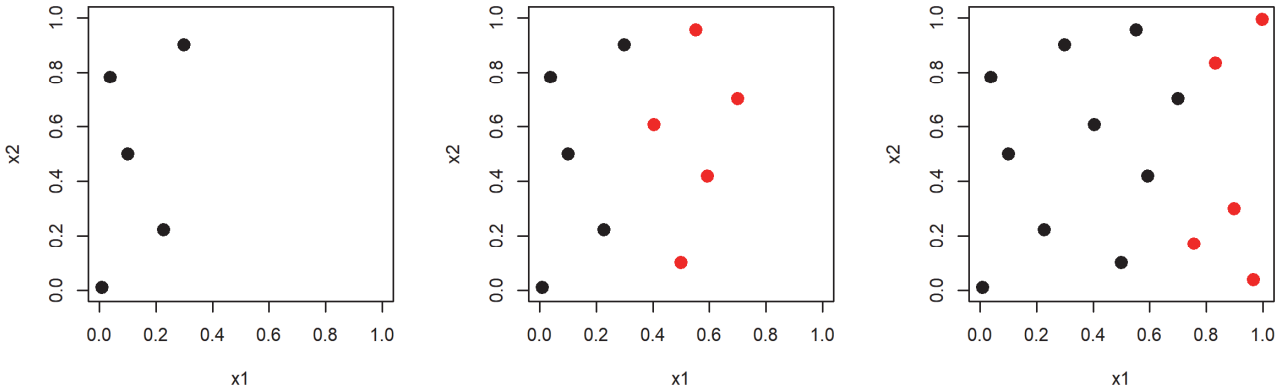
DOE: Design of Experiments

Goal: find designs that are space-filling at every stage of sequential design

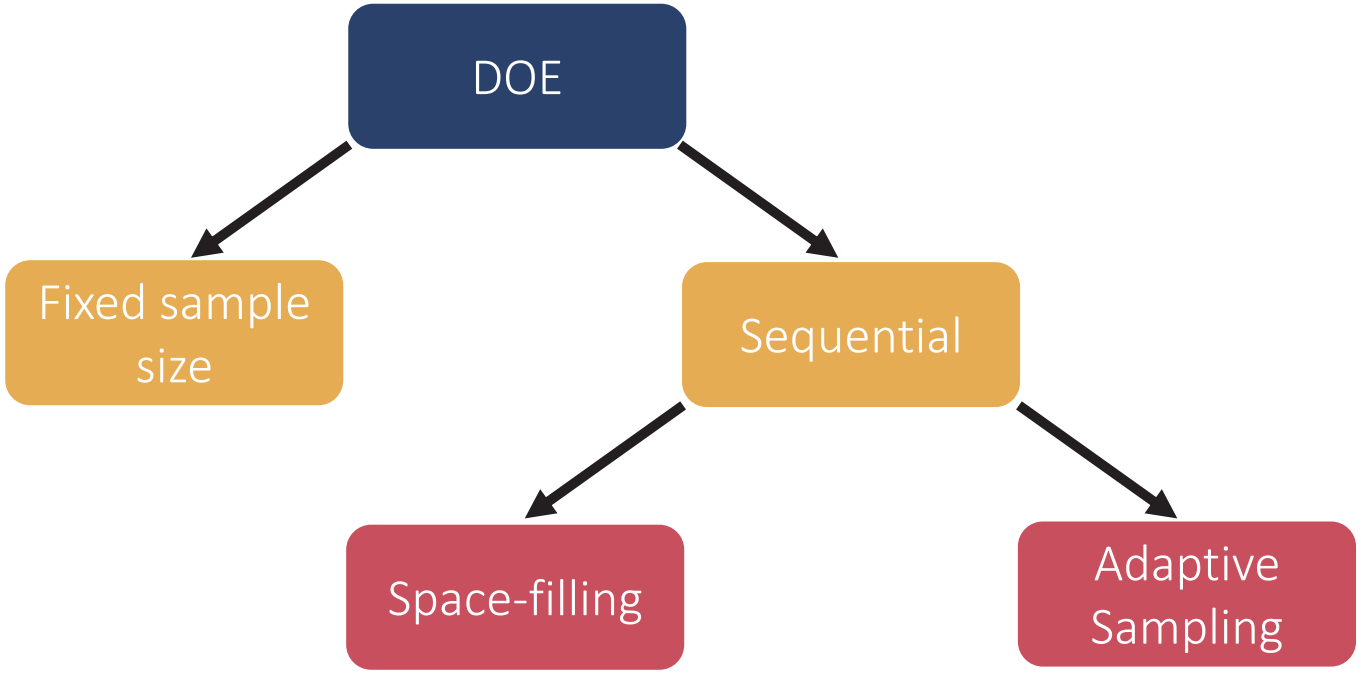
GOOD sequential design



BAD sequential design



Sequential test points can be found in two ways

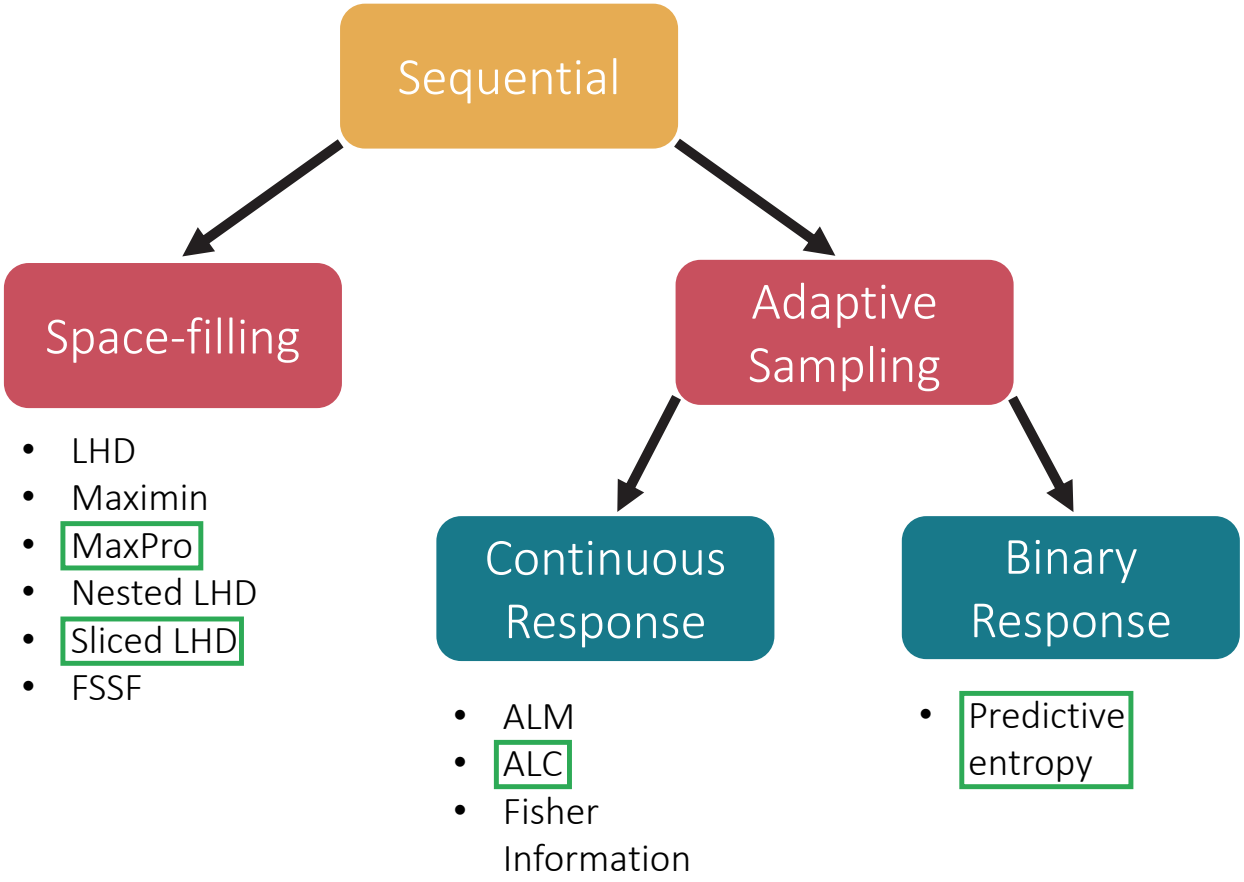


- Selected based on space-filling criterion
- Aims for coverage of operational space

- Selected based on model output
- Aims to cover spaces where predictive variance is high

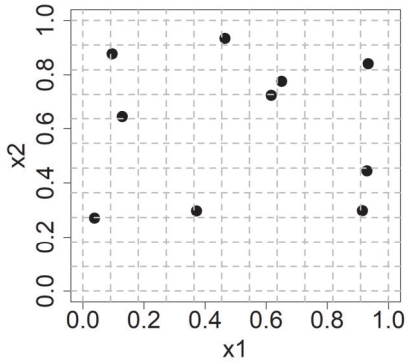
DOE: Design of Experiments

Available Sequential Space Filling Designs

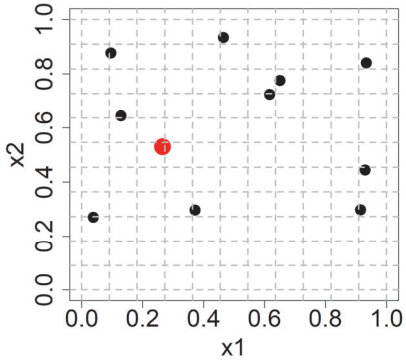


LHD: Latin Hypercube Design; MaxPro: Maximum Projection; FSSF: Fully Sequential Space-Filling; ALM: Active Learning MacKay; ALC: Active Learning Cohn

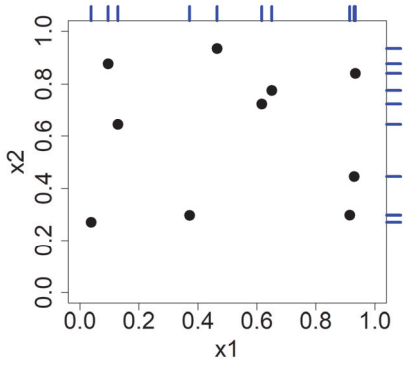
Latin Hypercube Design



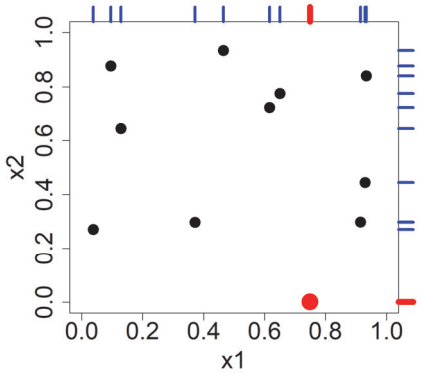
Place new test point in an unoccupied row and column



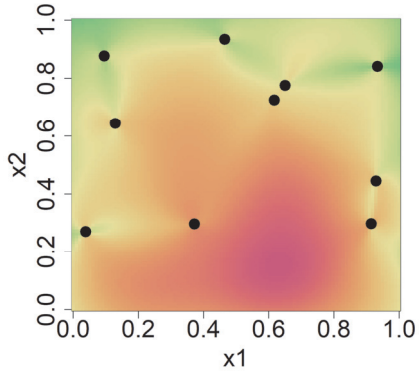
Maximum Projection



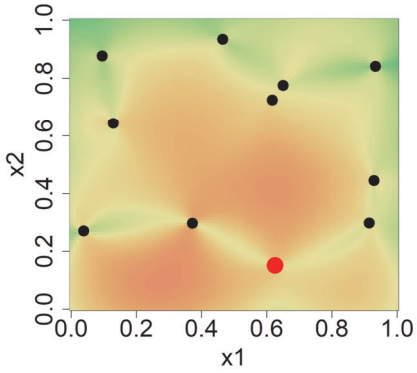
Place new test point to maximize projected distance



Adaptive Sampling

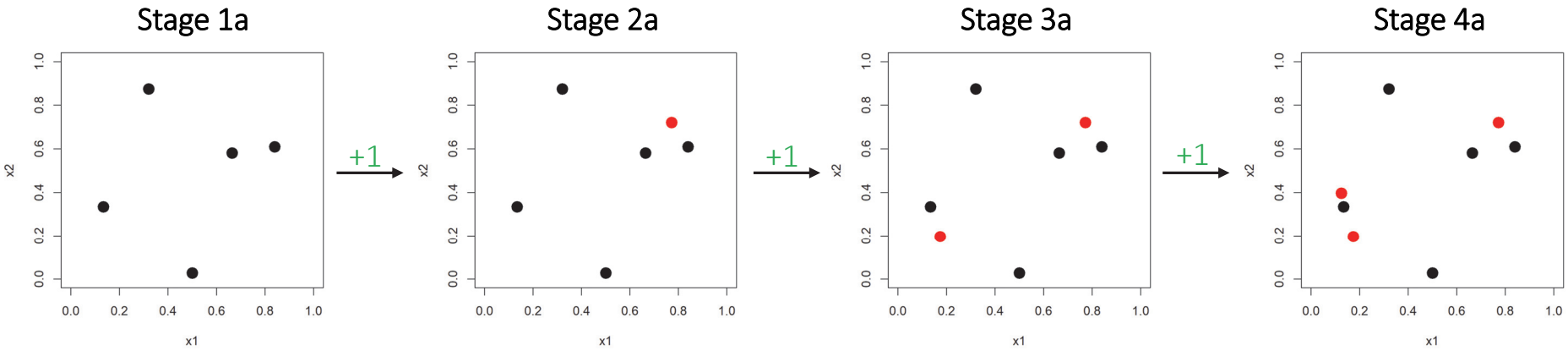


Place new test point to decrease predictive variance

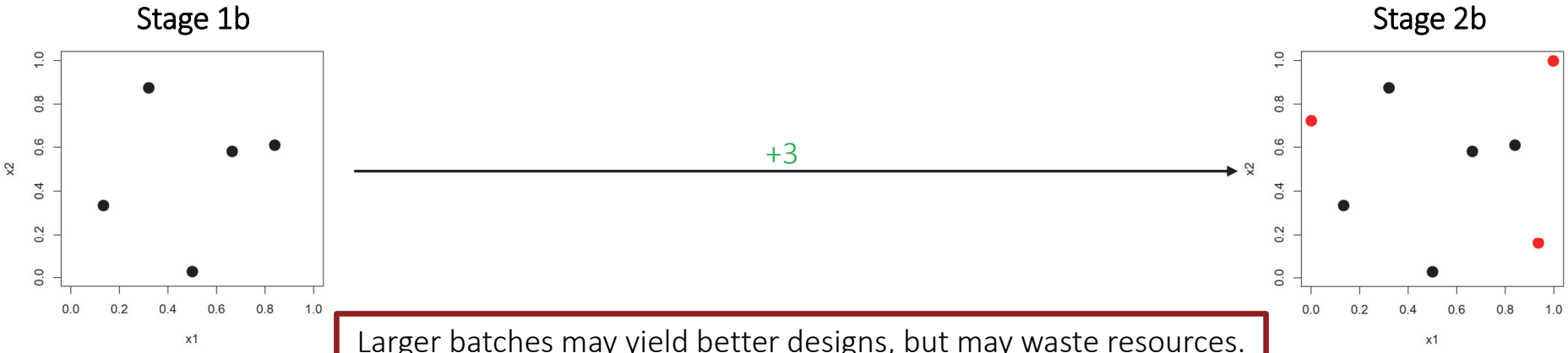


New test points can be added one at a time or in batches

One at a time (+1)



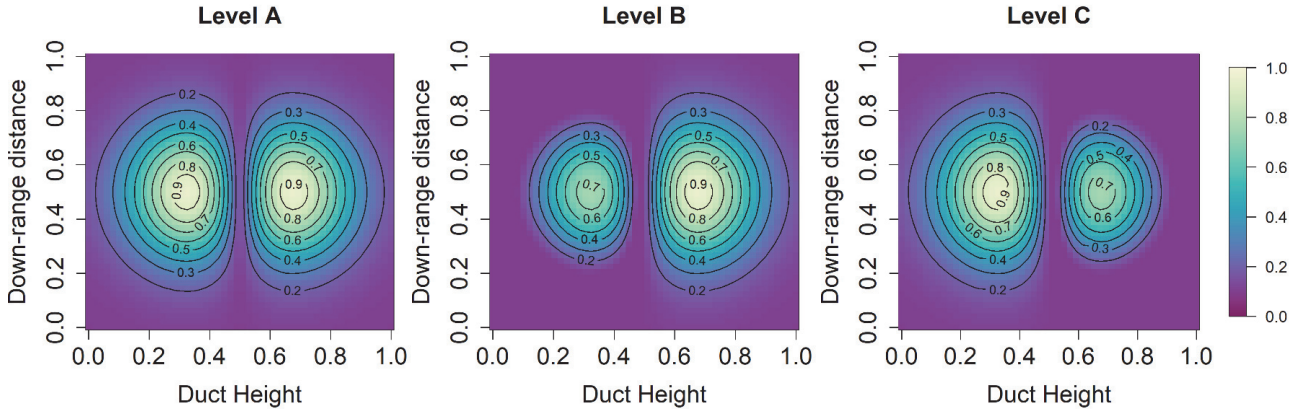
Batch (+3)



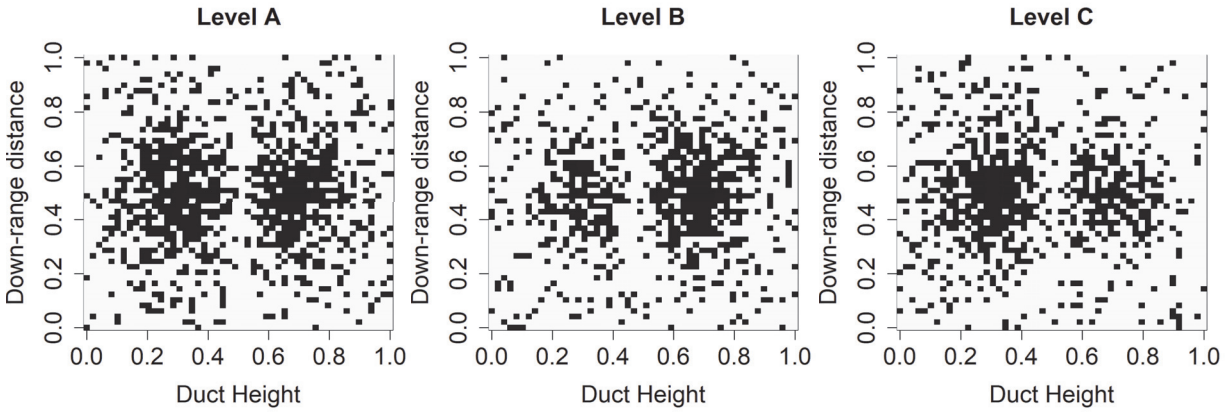
Larger batches may yield better designs, but may waste resources.

Visualization of data used in simulation*

Probability of Kill

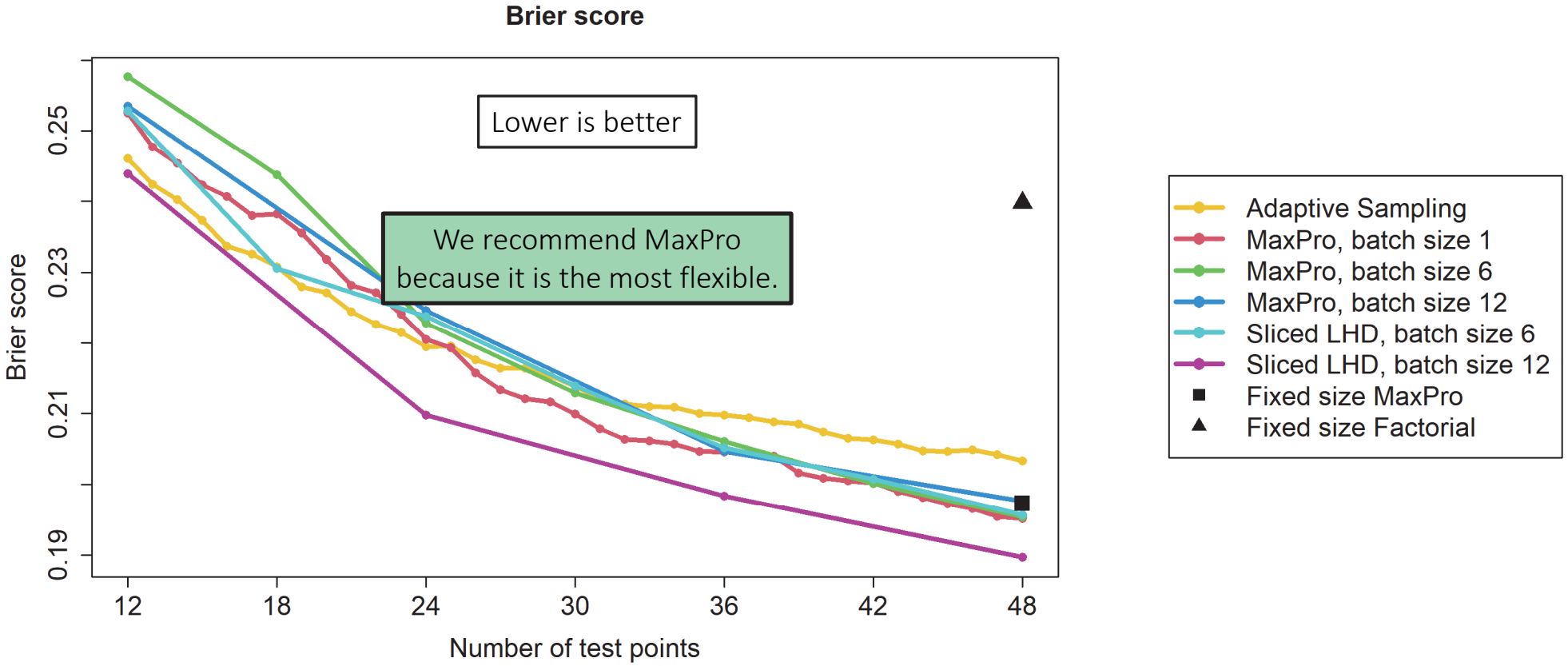


Kill/No-Kill



*All data are notional

Model performance at maximum sample size is roughly equivalent to fixed sample size performance



MaxPro: Maximum Projection; LHD: Latin Hypercube Design

Summary and Recommendations

- Sequential design does not require researchers to set sample size in advance
  - This can lead to greater test efficiency (e.g., saving time, money, and resources)
- It is important that designs be space-filling at every stage of sequential design process
- My project was to find sequential space-filling designs that were most applicable to IDA's work: those with existing software, and those which can handle categorical variables
  - Three methods fit these criteria: MaxPro, SLHD, and Adaptive Sampling
  - Since there is no clear winner, we recommend the MaxPro design for its flexibility
- The Test Science group should use these findings to coordinate with other research staff about when to use sequential DOE

MaxPro: Maximum Projection; SLHD: Sliced Latin Hypercube Design; DOE: Design of Experiments

Future Work

- We are writing a paper to document our findings for the T&E community

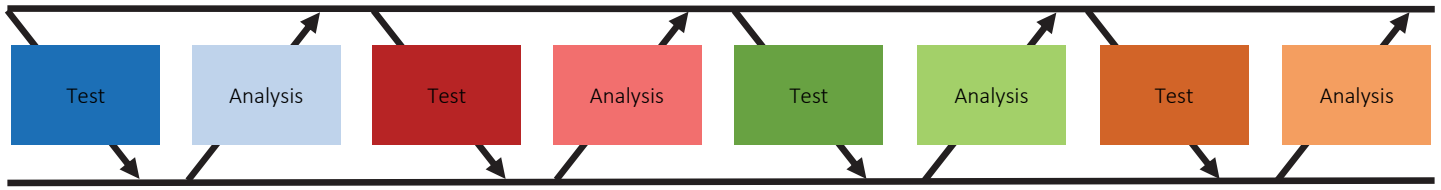
- Expand the research to metrics that indicate **when** to stop testing
  - The metrics used in the simulation would not be accessible during real testing

Thank you for listening! Questions?

Backup Slides

I conducted a simulation study to compare sequential space-filling methods

- Methods:
  - The three previously mentioned methods
  - Comparison: Fixed sample size designs, where size is maximum size of experiment
- Batch sizes:
  - 1, 6, 12
- Each rep begins with 12 test points, and then sequentially adds test points up to size 48
- Each test point is sampled 100 times to approximate kill/no-kill rate



Graphic taken from Test Design & Reporting course

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
20241105	OED Final		

4. TITLE AND SUBTITLE

Sequential Space-Filling Designs for Modeling & Simulation Analyses

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
HQ003424D0020		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
C9082		

6. AUTHOR(S)

Anna R. Flowers

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Institute for Defense Analyses
730 East Glebe Road
Alexandria, Virginia 22305

8. PERFORMING ORGANIZATION REPORT NUMBER

3003752

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Director, Operational Test and Evaluation
1700 Defense Pentagon
Washington, DC 20301

10. SPONSOR/MONITOR'S ACRONYM(S)

DOT&E

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution Statement A. Approved for public release: distribution is unlimited.

13. SUPPLEMENTARY NOTES

Project Leader: John Haman, jhaman@ida.org

14. ABSTRACT

Space-filling designs (SFDs) are a rigorous method for designing modeling and simulation (M&S) studies. However, they are hindered by their requirement to choose the final sample size prior to testing. Sequential designs are an alternative that can increase test efficiency by testing small amounts of data at a time. We have conducted a literature review of existing sequential space-filling designs and found the methods most applicable to the test and evaluation (T&E) community. We compare these methods using a notional example of the Evolved SEASPARROW Missile, Block 2.

15. SUBJECT TERMS

space-filling design, SFD, modeling and simulation, M&S, sequential design, test and evaluation, T&E

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
U	U	U	none	24

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
John Haman	(703) 845-2132

PREVIOUS EDITION IS OBSOLETE.

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18