INSTITUTE FOR DEFENSE ANALYSES

DATAWorks 2023: Implementing Fast Flexible Space-Filling Designs in R

Rebecca M. Medlin, Project Leader

Christopher T. Dimapasok
John T. Haman
Keyla Pagán-Rivera

April 2023

IDA Document NS D-33393

Log: H 2023-000042

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 contract HQ0034-19-D-0001, Task 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. V. Bram Lillard and consisted of Dr. Rebecca M. Medlin and Dr. Kelly M. Avery from the Operational Evaluation Division

For more information:
Dr. Rebecca M. Medlin, Project Leader
rmedlin@ida.org • (703) 845-6731

Dr. V. Bram Lillard, Director, Operational Evaluation Division
vlillard@ida.org • (703) 845-2230

Rigorous Analysis | Trusted Expertise | Service to the Nation

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-33393

DATAWorks 2023: Implementing Fast Flexible Space-Filling Designs in R

Rebecca M. Medlin, Project Leader

Christopher T. Dimapasok
John T. Haman
Keyla Pagán-Rivera

Executive Summary

Modeling and simulation (M&S) is a critical component of testing and evaluating major weapon systems in the Department of Defense. When planning M&S, testers use experimental design techniques to determine how much and which types of data to collect. When running M&S, analysts can explore multiple experimental design methodologies to apply when determining the best conditions to test.

Sometimes, testers employ full-factorial designs when planning a test using M&S. However, when designing a test that involves M&S, testers can use space-filling designs (SFDs) to better spread out points across the operational space. Fast flexible space-filling designs (FFSFDs) are a type of SFD useful for M&S because they work well in design spaces with disallowed combinations (i.e., nonrectangular design spaces) and permit the inclusion of categorical factors. Both nonrectangular design spaces and categorical factors are recurring features in defense testing. An example of the use of an FFSFD is the planning of M&S for missile testing.

Guidance from the Deputy Secretary of Defense and the Director of Operational Test and Evaluation (DOT&E) encourages the use of open and interoperable software and recommends the use of SFDs.[1] This project addresses the directives of these two memoranda.

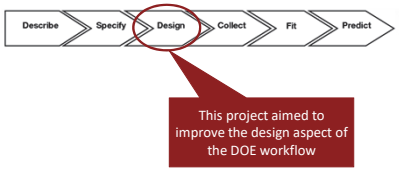
IDA analysts developed a function to create FFSFDs using the free statistical software R. To our knowledge, no R packages for creating an FFSFD can accommodate a variety of users' inputs, such as categorical factors. Moreover, users of IDA's function can share their code to make their work reproducible. IDA plans to make the function available through the Comprehensive R Archive Network (CRAN), which hosts other contributed R packages.

[1] See Deputy Secretary of Defense memorandum, May 2021, "Creating Data Advantage"; and Director of Operational Test and Evaluation memorandum, January 2017, "Clarifications on Guidance on the Validation of Models and Simulation Used in Operational Test and Live Fire Assessments."

Chris Dimapasok
Institute for Defense Analyses

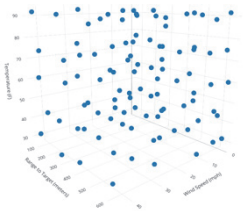
Implementing Fast Flexible Space-Filling Designs in R

Design of Experiments



What Is a Space-Filling Design?

Space-filling designs spread out points across the operational space. While space-filling designs are used for continuous data, fast flexible space-filling designs can incorporate **both continuous and categorical data**.



Project Goals

Space-filling designs are desirable for planning modeling and simulation test campaigns. No current R function provides space-filling designs that are suitable for categorical data. The main goal of this project was to develop a function in R that generates fast flexible space-filling designs and can:

- Handle continuous and categorical data
- Accommodate user input
- Be modified as an open-source R package

Run the R function

Set a seed for reproducibility

```
set.seed(730)
```

Generate a list of inputs

```
d <- list(wind_speed = c(1, 500),
          range_to_target = c(1,800),
          platform = c("A", "B", "C"))
```

Run the function

```
fff(d, n, r=10, N= NULL)
```

Input list

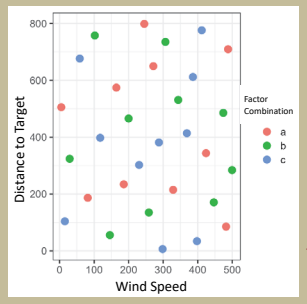
# of design points

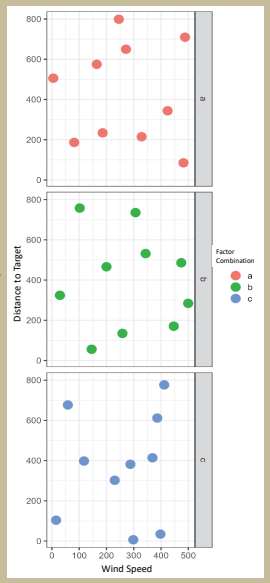
# of times labels are shuffled

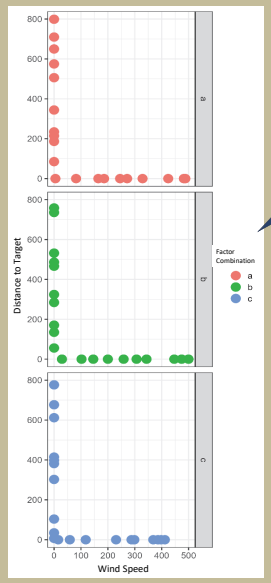
# of points to generate for clustering

The result is a space-filling design (SFD) that takes into account the categorical factors

Overall SFD with 2 continuous factors and 3 categorical factor combinations







The design has good projection properties

Mini SFD in each factor combination

Algorithm Workflow

Government Significance

The Deputy Secretary of Defense recommends open-source, easily accessible technology/interfaces. The Director of Operational Test and Evaluation (DOT&E) also emphasizes the importance of space-filling designs because they maximize opportunities to detect problems and can help quantify uncertainty in modeling and simulation.

Closing Remarks

- Designing experiments for modeling and simulation can be highly complex, especially if there are multiple factor types
- Fast flexible space-filling designs can be used when working simultaneously with continuous and categorical data
- Our R function generates fast flexible space-filling designs and can be used for future modeling and simulation work

Acknowledgments

Thank you to Keyla Pagán-Rivera, John Haman, Kelly Avery, and Rebecca Medlin for their review of this poster.

IDA

Implementing Fast Flexible Space-Filling Designs in R

Chris Dimapasok

Mentors: John Haman and Keyla Pagán-Rivera

OED

Institute for Defense Analyses

730 East Glebe Road ● Alexandria, Virginia 22305

Bottom Line Up Front

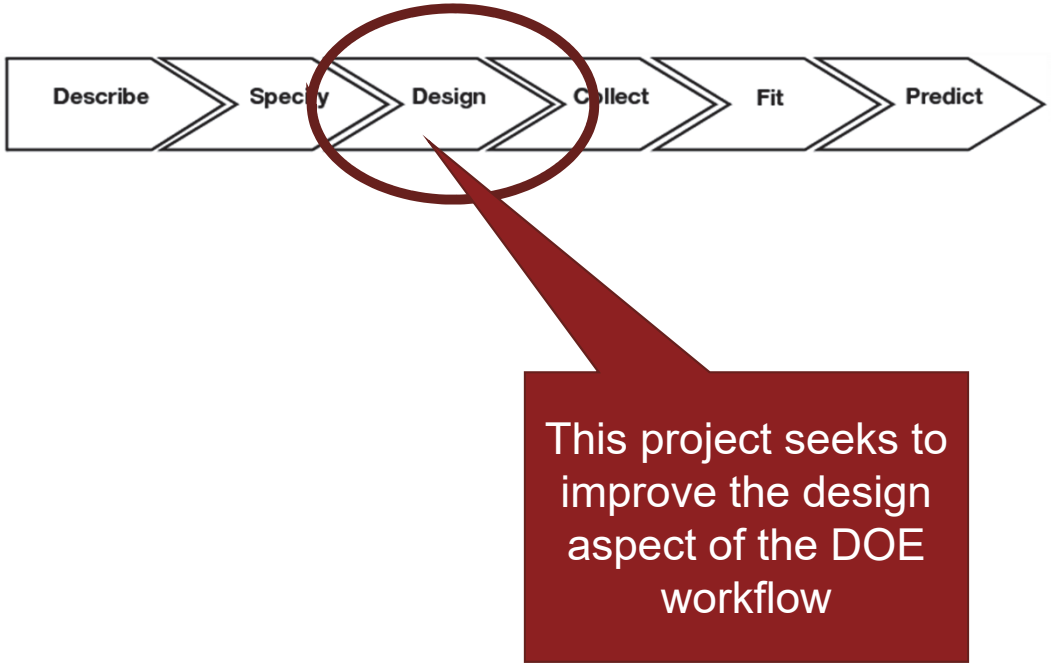
Motivation: Space-filling designs are desirable for planning modeling and simulation test campaigns

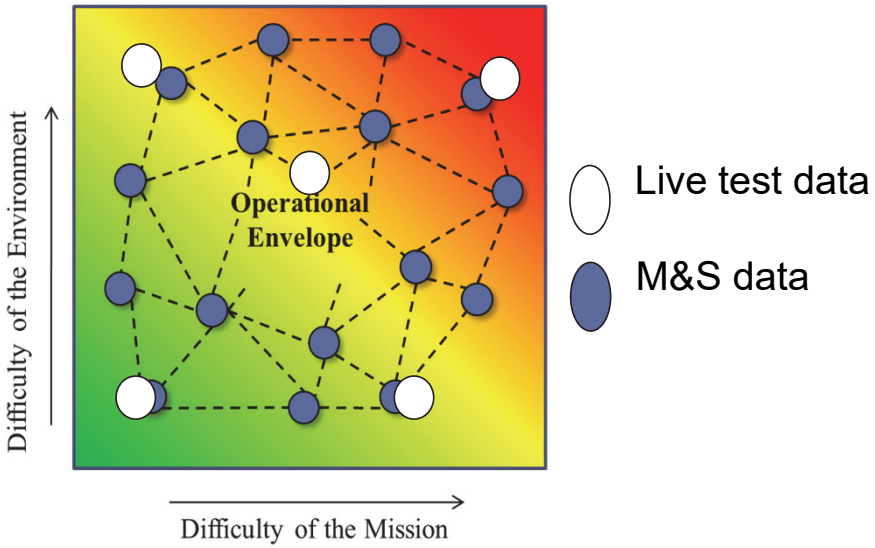
Problem: No R function provides space-filling designs that are suitable for categorical data

Project Goal: Develop a function in R that generates fast flexible space-filling designs and can:

- Handle continuous and categorical factors
- Accommodate user inputs
- Be edited (as an open-source R package)

IDA leverages DOE techniques to choose test conditions when planning a test for a military system



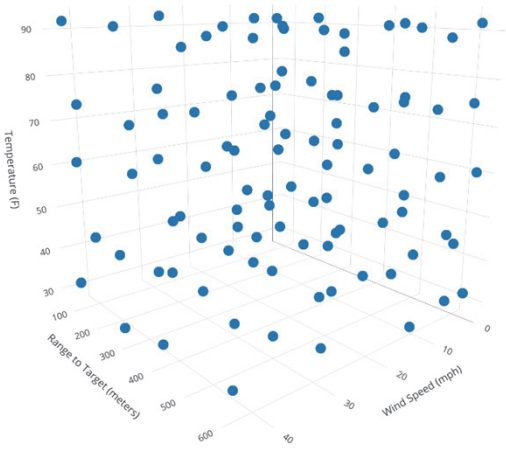


DOE = design of experiment

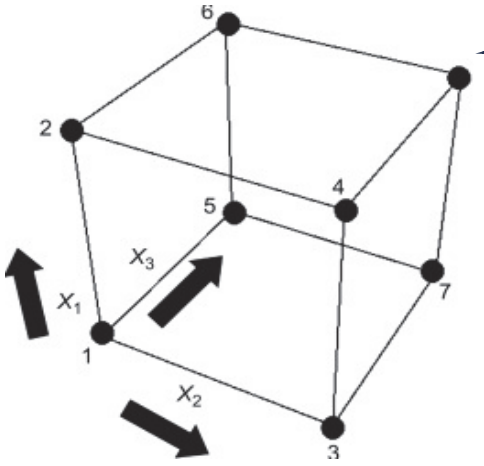
M&S = modeling and simulation

Overview of Space-Filling Designs

- Space-filling designs are used to spread out points across the operational space
- Space-filling designs differ from classical designs, which tend to push points to the exterior of the operational space



Space-Filling Design



Replicates

Factorial Design

M&S = modeling and simulation

Our R program creates an FFSFD, a type of space-filling design

- Operational test plans often require continuous and categorical variables
- An FFSFD can take categorical and continuous data as inputs and can accommodate nonrectangular design spaces
- Most other space-filling designs apply only to continuous factors

Categorical Data Examples	Continuous Data Examples
Rainy/sunny/cloudy	Temperature
Windy/not windy	Wind speed
Launch platform of a missile	Distance to target

FFSFD = fast flexible space-filling design

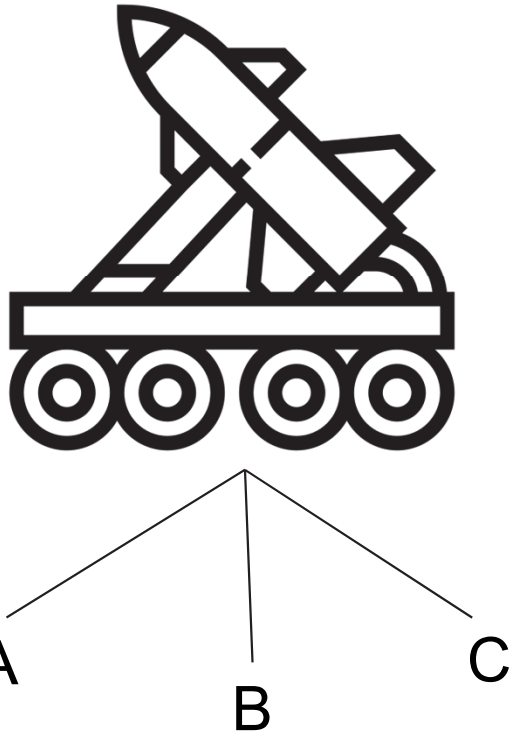
* The R program implements the algorithm Ryan Lekivetz and Bradley Jones propose in their 2019 paper, "Fast Flexible Space-Filling Designs with Nominal Factors for Nonrectangular Regions."

Implementation of FFSFD: Missile System Example

Wind Speed
(Continuous)

Distance to Target
(Continuous)

Launch Platform
(Categorical)



Response: Missile Accuracy (Miss Distance)

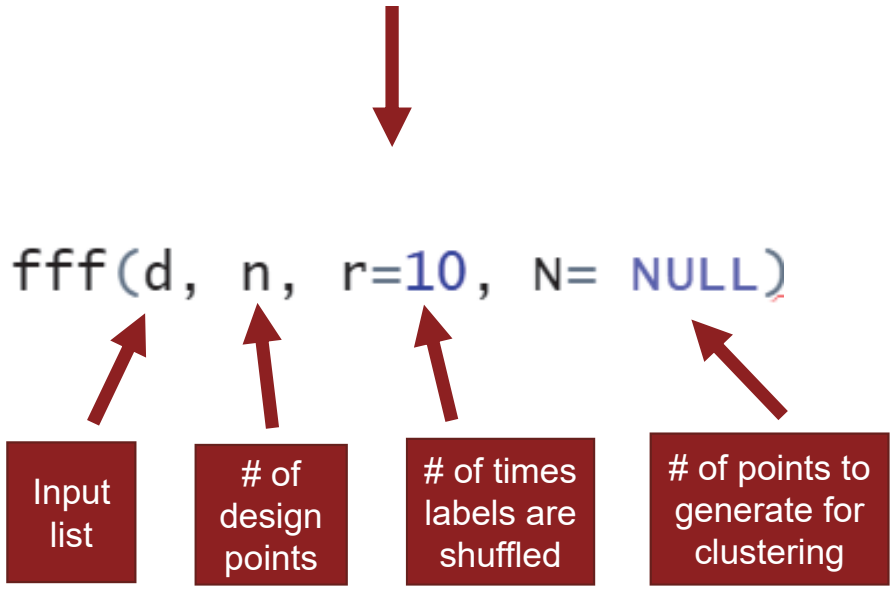
FFSFD = fast flexible space-filling design

Icon source: NounProject.com

This R package is free, open-source software

```
d <- list(wind_speed = c(1, 500),
          range_to_target = c(1,800),
          platform = c("A", "B", "C"))
```

Step 1: Generate list of inputs

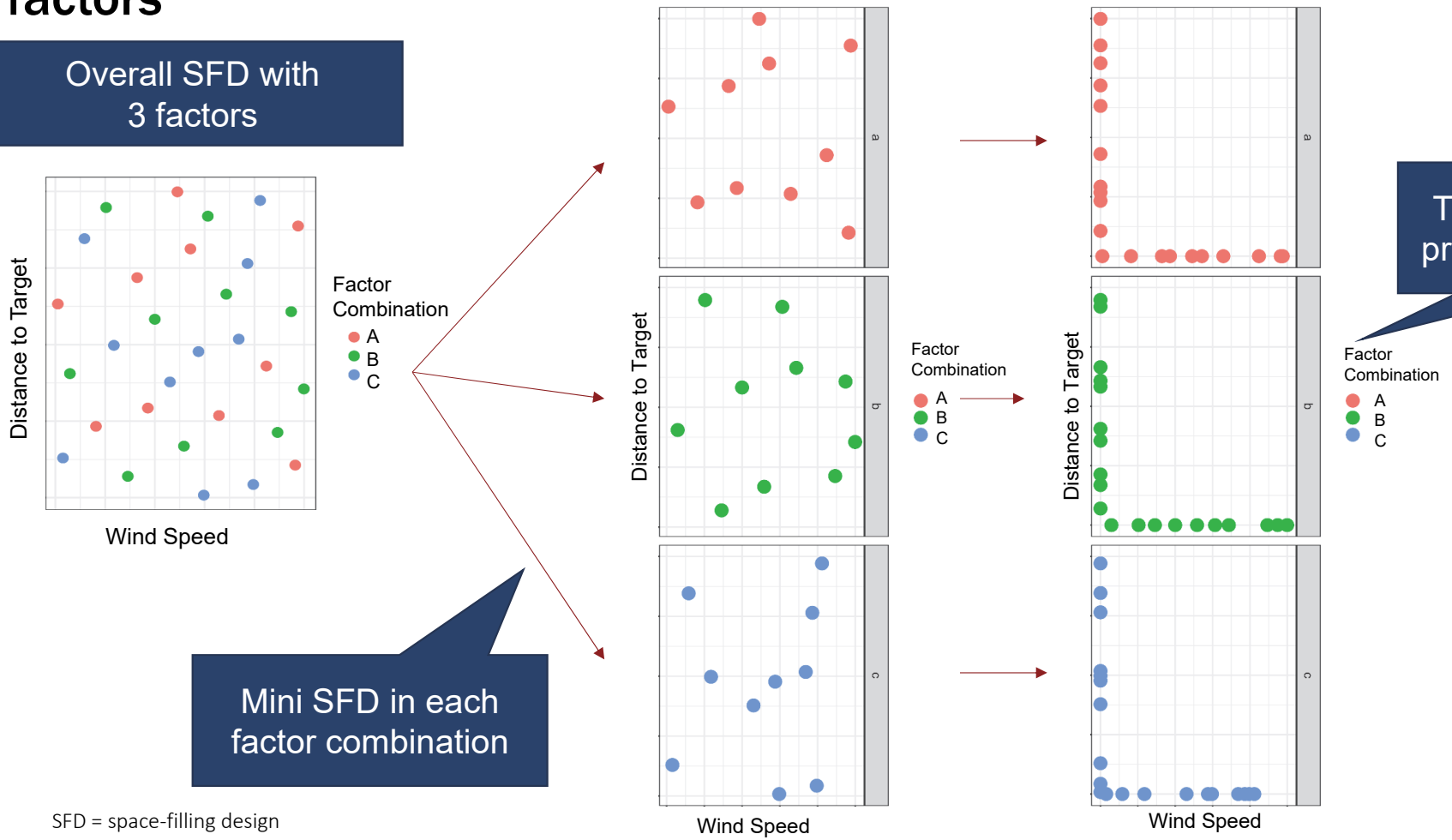


Step 2: Plug into fff function and generate design

Ryan Lekivetz and Bradley Jones, 2014, "Fast Flexible Space-Filling Designs for Nonrectangular Regions," *Quality and Reliability Engineering International* 31, no. 5, 829–837

The result is a space-filling design that takes into account categorical factors

Overall SFD with 3 factors



Mini SFD in each factor combination

The SFD has good projection properties

SFD = space-filling design

DEPSECDEF is pushing the Department toward more open and interoperable software. DOT&E recommends SFDs, but the most useful methods are proprietary and hard to reproduce.

DEPSECDEF's 2021 memo on Creating Data Advantage[1]

1. Maximize data sharing and rights for data use: all DoD data is an enterprise resource.
2. Publish data assets in the DoD federated data catalog along with common interface specifications.
3. Use automated data interfaces that are externally accessible and machine-readable; ensure interfaces use industry-standard, non-proprietary, preferably open-source, technologies, protocols, and payloads.
4. Store data in a manner that is platform and environment-agnostic, uncoupled from hardware or software dependencies.
5. Implement industry best practices for secure authentication, access management, encryption, monitoring, and protection of data at rest, in transit, and in use.

My project combines the directives of these two memos.

DOT&E's 2017 memo on M&S Validation[2]

- A robust design for the M&S that systematically covers the range of operationally realistic inputs over which the model will be accredited. Space-filling design methodologies are preferred because they not only maximize opportunities for problem detection, but also support the development of statistical emulators that can be compared to live data and assist in quantifying uncertainty in the M&S.

DEPSECDEF = Deputy Secretary of Defense; DOT&E = Director of Operational Test and Evaluation; M&S = modeling and simulation; SFD = space-filling design

[1] Deputy Secretary of Defense memorandum, May 2021, "Creating Data Advantage."

[2] Director of Operational Test and Evaluation memorandum, January 2017, "Clarifications on Guidance on the Validation of Models and Simulation Used in Operational Test and Live Fire Assessments."

Closing Remarks

- Designing experiments for modeling and simulation can be highly complex, especially if there are multiple factor types
- Fast flexible space-filling designs can be used when working simultaneously with continuous and categorical data
- Our R package generates fast flexible space-filling designs and can be used for future modeling and simulation work
- Coming soon: Function will be published as an R package on CRAN and testscience.org

Acknowledgments

- John Haman and Keyla Pagán-Rivera
- Jason Sheldon
- Kelly Avery
- Rebecca Medlin
- Curtis Miller
- Mark Herrera
- Joanna “Jojo” C. Di Scipio
- Miriam Armstrong
- Lamarr Colvin, Heather Neff, and HR Team
- Martha Sherell Ard Smith
- Kerri A. Mahalla
- OED

Backup

Common questions on fast flexible designs and space-filling designs

Q: Why not simply generate random points in the design space?

A: Random designs do not actually have good space-filling properties. They tend to be too clumpy and have gaps.

Q: Why use an inelegant clustering algorithm?

A: It is awkward, but we benefit from being able to generate designs on nonrectangular design spaces.

Q: Does this create structure from nothing?

A: No. We assume that the M&S is low-noise or noiseless, and these designs give us the best chance to model that kind of response.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
26-04-2023	Final	

4. TITLE AND SUBTITLE

DATAWorks2023 Implementing Fast Flexible Space Filling Designs In R

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mr. Christopher T. Dimapasok
Dr. John T. Haman
Dr. Keyla Pagan-Rivera

5d. PROJECT NUMBER

C9082

5e. TASK NUMBER

C9082

5f. WORK UNIT NUMBER

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

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

8. PERFORMING ORGANIZATION REPORT NUMBER

NS D-33393

H 2023-000042

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

10. SPONSOR/MONITOR'S ACRONYM(S)

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

12. DISTRIBUTION/AVAILABILITY STATEMENT

This OED Draft has not been approved by the sponsor for distribution and release. Reproduction or use of this material is not authorized without prior permission from the responsible IDA Division Director.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Modeling and simulation (M&S) can be a useful tool for testers and evaluators when they need to augment the data collected during a test event. During the planning phase, testers use experimental design techniques to determine how much and which data to collect. When designing a test that involves M&S, testers can use Space-Filling Designs (SFD) to spread out points across the operational space. Fast Flexible Space-Filling Designs (FFSFD) are a type of SFD that are useful for M&S because they work well in nonrectangular design spaces and allow for the inclusion of categorical factors. Both of these are recurring features in defense testing. Guidance from the Deputy Secretary of Defense and the Director of Operational Test and Evaluation encourages the use of open and interoperable software and recommends the use of SFD. This project aims to address those.

IDA analysts developed a function to create FFSFD using the free statistical software R. To our knowledge, there are no R packages for the creation of an FFSFD that could accommodate a variety of user inputs, such as categorical factors. Moreover, by using this function, users can share their code to make their work reproducible.

15. SUBJECT TERMS

Design of Experiments (DOE); Modeling and Simulation (M&S); Space Filling Designs

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Rebecca M. Medlin
Unclassified	Unclassified	Unclassified	SAR		19b. TELEPHONE NUMBER *(Include area code)* 703-845-6731