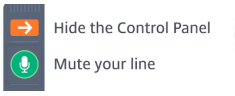
Space-Filling Designs and Gaussian Process Models

STAT-EASE 360

Martin Bezener, PhD
President & Chief Technology Officer
martin@statease.com

February 2023

Making the most of this learning opportunity



To prevent audio disruptions, all attendees will be muted.

Questions can be posted in the **Question** area. If they are not addressed during the webinar, I will reply via email afterwards.

Questions may also be sent to stathelp@statease.com. Please provide your company name and, if you are using Design-Expert, the serial number (found under Help, About).

Note: The slides and a recording of this webinar will be posted on the Webinars page of the Stat-Ease website within a few days.

2

Agenda

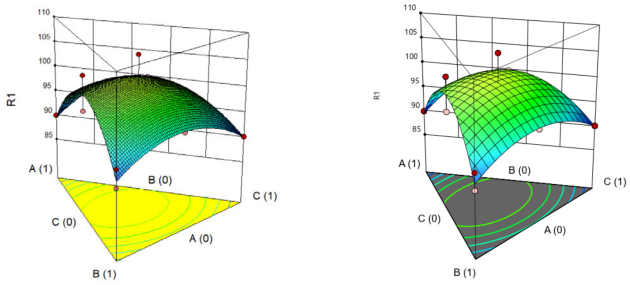
- Introduction to Stat-Ease 360
- Space-Filling Designs
- Gaussian Process Models
- Demos and Conclusion

Agenda

- **Introduction to Stat-Ease 360**
- Space-Filling Designs
- Gaussian Process Models
- Demos and Conclusion

Design-Expert

- Design-Expert (DX) has been a leading software package in the design and analysis of experiments (DOE).
- Originally developed in the mid-1980s by Pat Whitcomb, DX provides cutting-edge, easy-to-use tools that don't require lengthy, formal statistical training.



Design-Expert – Going Forward

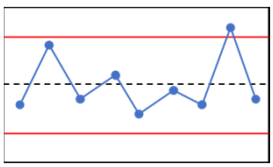
- Over the years, Stat-Ease has received steady requests for more advanced DOE tools.
- We've also received demand for tools that are adjacent to 'old school' DOE, such as **design and analysis of computer experiments**, Gage R&R (measurement systems analysis), more computing capabilities, statistical process control, and many more.
- In the interest of keeping Design-Expert as simple and accessible as possible, we have historically tabled these types of requests.
- The demand was simply too great!

STAT-EASE 360

- Stat-Ease 360 (SE360) was released **October 4, 2021**.
- It contains a **full copy** of Design-Expert software.
- The first release of SE360 **also** contained some of our most highly-requested advanced DOE features:
  - **Space-filling designs**
  - **Gaussian process models for zero-error data**
  - Python scripting
  - Advanced classification tools for logistic regression
- Recently we also added:
  - Excel Import
  - Random Blocks

STAT-EASE 360

- Going forward, Design-Expert will continue to be targeted towards engineers, formulators, and others working in R&D. Remember, SE360 contains **all of DX**.
- Going forward, easy-to-use and bread-and-butter techniques will go into both DX and SE360.
- Highly-advanced, non-DOE, and highly technical & computational features will only go into SE360. Some ideas we have for the future:
  - Measurement Systems Analysis
  - Multivariate Analysis
  - Advanced Scripting/Computing



Agenda

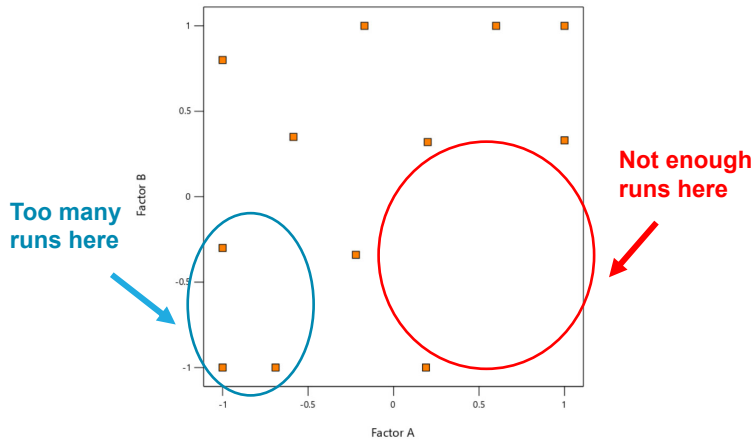
- Introduction to Stat-Ease 360
- **Space-Filling Designs**
- Gaussian Process Models
- Demos and Conclusion

Space-Filling Designs

- Space-Filling Designs (SFDs) are a category of experimental designs that aim to 'best' cover or spread out runs in the experimental design space.
- We generally don't want large gaps in the design space that are devoid of runs, and we also don't want too many points clustered in the same region.
- While SFDs are commonly associated with computer experiments, they also have uses in physical (traditional) experiments.
- SFDs are still a very active research area – there are many open questions. It's not always obvious what the 'best' space-filling design will be for a given experimental situation.

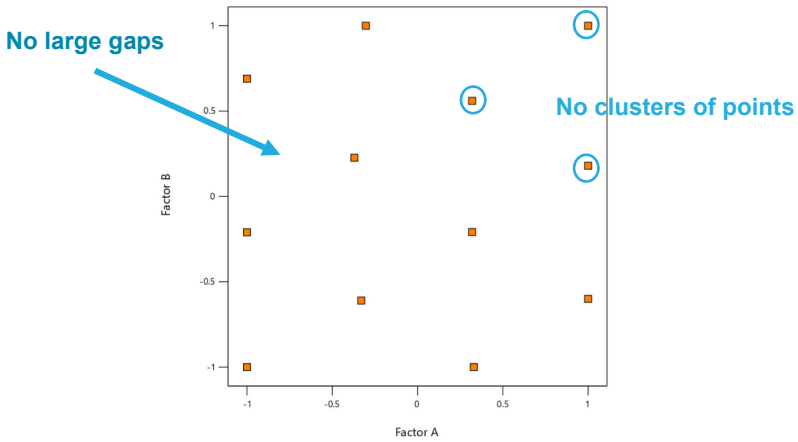
Space-Filling Designs

Here's a simple example of a design that's **NOT** space-filling:



Space-Filling Designs

Here's a simple example of a design that **IS** space-filling:



Space-Filling Designs

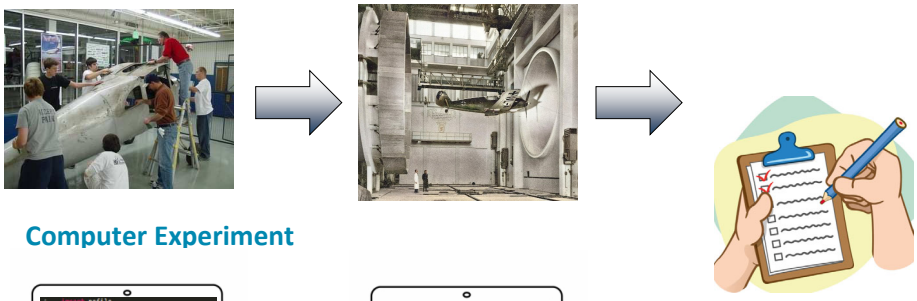
- We'll now cover a few different questions:
  - What is computer/simulation experiment?
  - What are the major classes of space-filling designs?
  - Can space-filling designs be used in physical experiments?
  - Strategies and tips?
  - How to use SFDs in Stat-Ease 360?

What is a Computer Experiment?

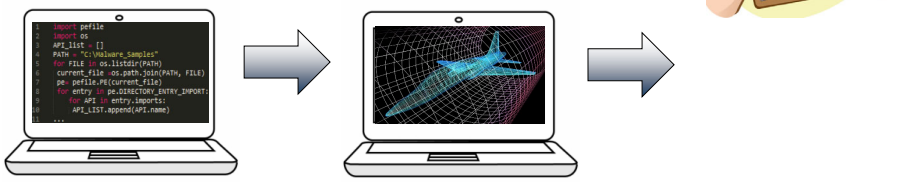
- Computer experiments can differ from physical experiments in the following ways:
  - For a given set of factors, the output is deterministic (no error)
  - **A computer simulation, rather than a physical experiment in the lab, is run to obtain the output.**
  - Repeating a simulation for a given set of factors will usually produce identical results.
  - Each run can be **extremely** time-intensive to obtain.
- Because there is no error, certain desirable features of experimental designs, like replicates, no longer make sense.

What is a Computer Experiment?

Physical Experiment



Computer Experiment

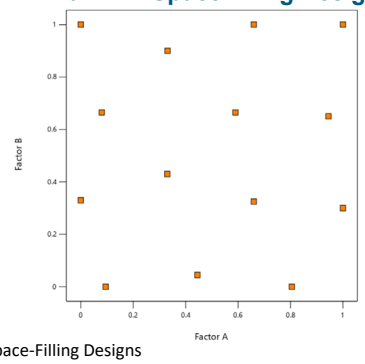


Space-Filling Designs 15

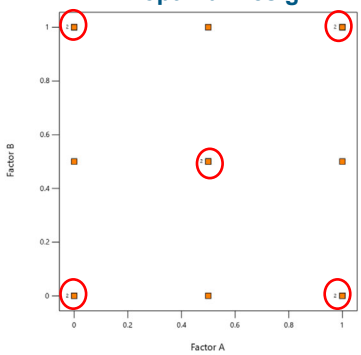
What is a Computer Experiment?

- Because the output has no error, replicates don't make sense.
- Designs for computer experiment are typically space-filling – that is, they aim to cover the design space without leaving large gaps between design points.

Maximin Space-filling Design



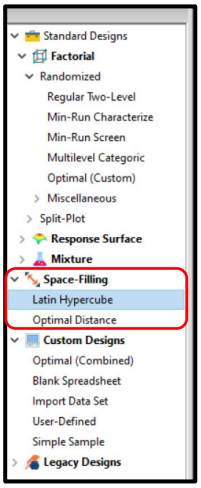
D-optimal Design



Space-Filling Designs 16

Design and Analysis of Computer Experiments

- **Demo:** Space-Filling Designs in **SE360**:



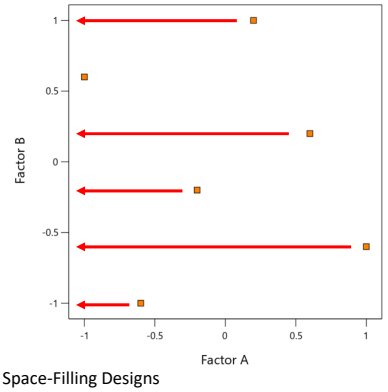
Major Classes of Space Filling Designs

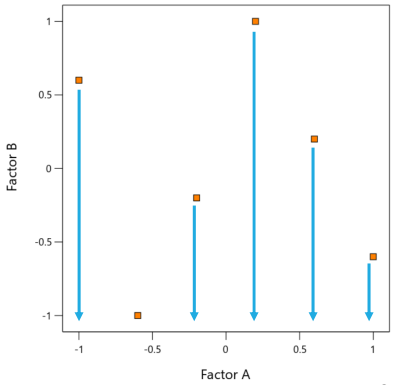
- Stat-Ease 360 has two major offerings for space-filling designs:
  - **Latin Hypercube Designs**
  - **Optimal Space-Filling Designs**
    - Model-based
    - Distance-based

Major Classes of Space Filling Designs

Latin Hypercube Designs (LHDs):

- Start with **k** factors. Divide each factor into **p** levels.
- Each factor takes each level exactly one time in the design:

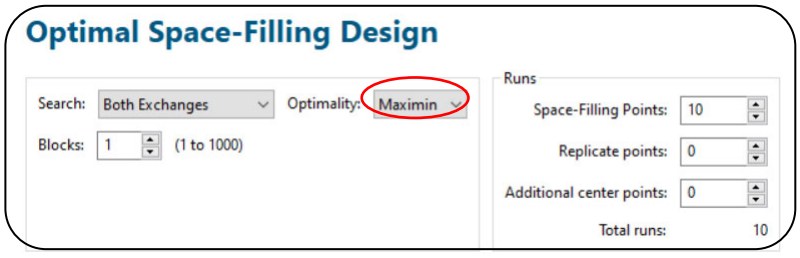




Major Classes of Space Filling Designs

Optimal Space-Filling Designs:

- These are generally built the same way as the 'usual' optimal designs (**D**, **I**, etc.).
- The current criteria for building these is Maximin – which tries to build the design that maximizes the minimum distance between runs.
- The UI has been slightly tweaked for the space-filling case:



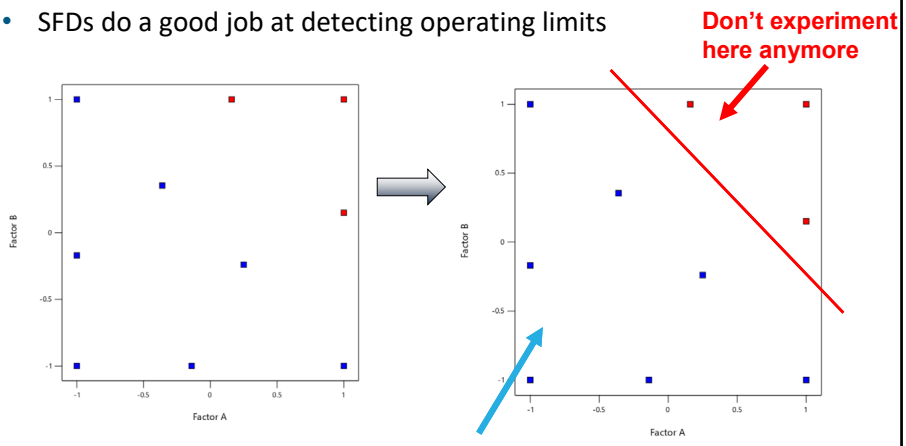
Space-Filling Designs for Physical Experiments

- Space-filling designs can **also** be used in physical experiments:
  - When the experimental error either (1) known or (2) very low, replicates may not be of much value. In this case, a space-filling design could be useful.
  - If you are performing an exploratory study on a new design space, a space-filling design will generally test many unique points that are spread far apart. This can help eliminate unfeasible runs and further refine the design space.
  - If you expect the response to be very curvy or spiky in a small region of the design space, then a space-filling design will have a better chance of detecting that behavior.

Space-Filling Designs for Physical Experiments

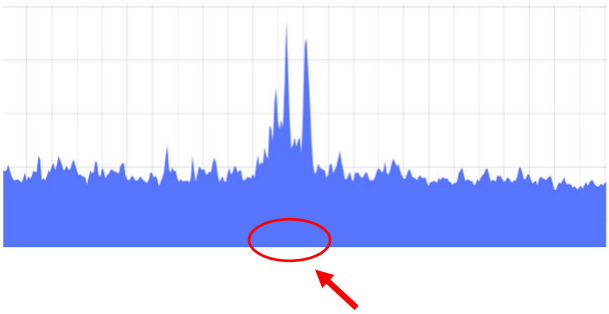
Exploratory Experimentation

- SFDs do a good job at detecting operating limits



Space-Filling Designs for Physical Experiments

- If the response you are measuring is 'spiky' or has a small region of unusual behavior, a space-filling design will be more likely to detect this area of the curve.



Agenda

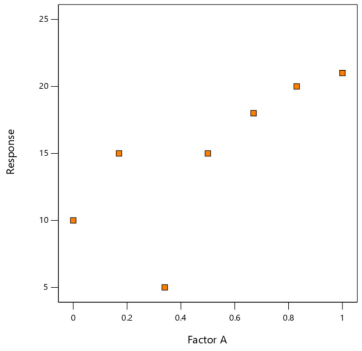
- Introduction to Stat-Ease 360
- Space-Filling Designs
- **Gaussian Process Models**
- Demos and Conclusion

Gaussian Process Models

- **Next Question:** how do we analyze this type of data?
- If a SFD is used for a physical experiment, the usual methods like OLS regression can be used. This is the easy case.
- Ordinary polynomial models are often not adequate for modelling data that comes from a computer experiment, however, because data from a computer experiment has no experimental error.

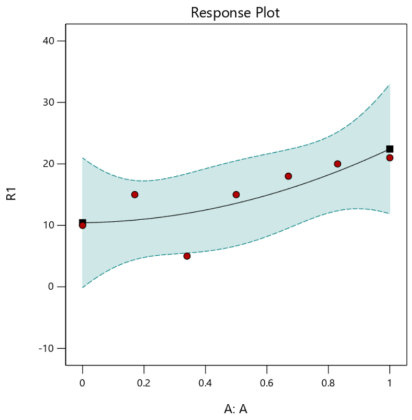
Design and Analysis of Computer Experiments

- Suppose we collect data from a simulation experiment that looks like this:



Design and Analysis of Computer Experiments

- A polynomial OLS model would look like this:



Gaussian Process Models

- The big issue is that the model should predict the response **perfectly**, if there is no error!
- It is impossible to do this with a polynomial model, unless the model has as many terms as there are runs in the data set.
- Gaussian Process Models (GPMs) are one tool that are used to model zero-error data obtained from a computer experiment.
- GPMs are best illustrated with a demo.

Agenda

- Introduction to Stat-Ease 360
- Space-Filling Designs
- Gaussian Process Models
- **Demos and Conclusion**

Demo #1

- The first demo is a one-factor experiment. The purpose is to show the SE360 user-interface, and to show how Gaussian Process Models work.
- The is one factor, which is **time** which goes from 0 to 100 minutes. The response is obtained via simulation. Repeating a simulation at the same **time** will produce the same exact output, so replicates are not necessary.
- The simulated response was the expected **output** of a process.
- There was enough time for a total of 6 simulations.

Demo #2

- The second demo is an example from the medical device industry. Unfortunately, the factors/responses had to be somewhat obfuscated for confidentiality.
- The R&D team was developing a new artificial joint. They wanted to see how the stiffness of the joint changed as two factors were varied.
- It is very difficult to perform this as a physical experiment. So as a first step, the team used a computer experiment.
- They were able to create a simulation that described how the stiffness of the new joint was expected to change as the two factors were tweaked. The simulation was very time consuming (1 run took about 40 hours to obtain).
- There was enough time for 8 runs.

Conclusion

- Space-Filling Designs are an interesting class of DOEs that have many practical applications.
- While they are primarily associated with computer experiments, they can also be used in physical experiments as well.
- SE360 offers a solid foundation for experimenters wishing to use space-filling designs in their own work.
- More features will be released soon:
  - SFDs for mixture experiments
  - Design that incorporate categorical and numerical factors
  - Additional GPM tools and statistics

Python Scripting Capabilities

- SE360 also includes the ability to write Python scripts. Here is just a small sampling of what you can do:
  - Write scripts to automate routine processes
  - Create simulations
  - Combine features of Stat-Ease 360 with features from relevant Python packages (e.g. scikit-learn)
  - Create infinitely customizable plots and graphs
  - Facilitate import/export of data between Design-Expert and other software

Python Scripting Capabilities