

# Space-Filling Designs for Multi-Layer Nested Factors

Ryan W. Gryder<sup>1</sup>, Sara R. Wilson<sup>2</sup>, Kurt A. Swieringa<sup>2</sup>, and David J. Edwards<sup>3</sup>

<sup>1</sup>Boston University, Boston, Massachusetts 02215

<sup>2</sup>National Aeronautics and Space Administration, Hampton, Virginia 23681

<sup>3</sup>Virginia Commonwealth University, Richmond, Virginia 23284

## Abstract

This article considers computer experiments where levels for continuous factors are selected in sequential order with the level selected for one factor directly affecting the range of possible levels for the nested factor, and so on for a finite number of factors. In addition, we assume the nested relationships between the factors have no closed form solution. In this paper, we propose an approach for constructing a multi-layer nested factor design, or multi-NFD for short. This space-filling design approach takes advantage of the maximin criterion and can be analyzed using a standard Gaussian process model. While the multi-NFD approach can be adapted for future computer experiments involving factor relationships of this type, we present results from a particular aerospace computer simulation study.

**Key Words:** branching factor, computer experiment, Gaussian process, maximin criterion, simulation.

## Introduction

In this article, we define multi-layer nested factors as those that are continuous and ordered in their selection. That is, factor levels are selected sequentially with the level selected for

one factor affecting the range of possible values for a subsequent nested factor. Thus, each factor is nested within another factor, but the exact nested relationships have no closed form solution. In this paper, we describe our process of engineering an appropriate *space-filling design* for an aerospace computer simulation in order to appropriately fill the experimental design region using a finite number of points. Our pursuit of a space-filling design takes advantage of the *maximin criterion*, which is a common metric used to evaluate space-filling designs. We refer to the proposed design as a ***multi-layer nested factor design***, or ***multi-NFD***. While the unique multi-layered nested factor relationship originated from a very specific aerospace application, we emphasize that our proposed solution is a general approach that can be adjusted for any computer experiment involving a similar factor relationship. We also briefly discuss the analysis of such a design using *Gaussian process modeling* and some results from our specific application.

For statisticians and engineers, computer experiments have increased in popularity as a means of emulating physical experiments due to advancements in computational efficiency and storage. A computer experiment is executable computer code that mathematically models or emulates an actual physical experiment via codified inputs and outputs. Computer experiments are beneficial when the physical experiment is a strain on resources or researchers want to investigate new technologies in a simulation environment prior to implementation (Montgomery, 2012).

A stochastic computer experiment contains a random component with two runs of the same set of inputs resulting in possibly two different output responses, which implies a need for replicates. Conversely, a deterministic computer experiment has no random component, and so the same set of inputs will yield the exact same outputs if the run is repeated. The application described in this paper focuses on the deterministic case. Contrary to physical experiments, one does not need to consider the usual elements of experimental design such as randomization, replication, and blocking in a computer experiment where a random component is absent. Additionally, all variation in output depends solely on the different

treatment combinations. A common strategy to select the factor levels for a computer experiment is to use a *space-filling design*, which strategically selects a small number of uniformly distributed points to fill the experimental region. We strive for uniformity in the sampling of design points since predictive models of space-filling designs are usually based on interpolating a response surface over the design region. Unevenly spaced design points can lead to poor predictions in the sparsely sampled portions of the design space (Dean et al., 2015).

A common space-filling design for continuous factors is a *Latin Hypercube Design*, or LHD, first introduced by McKay et al. (1979). These designs are generalizations of the Latin Square Design in which there is one run per row and column combination of each two-factor projection and have the desired property of one-dimensional uniformity. Modifications to the classic LHD have been proposed through use of the *maximin criterion* (Johnson et al., 1990). This criterion seeks a design that maximizes the minimum distance between every pair of design points. If there are  $m$  continuous factors that make up a design region  $A \subset \mathbb{R}^m$ , then the maximin criterion seeks a space-filling design that maximizes the following:

$$\min_{a_i, a_j \in A; i \neq j} d(a_i, a_j) \quad (1)$$

where  $d$  is some measure of the distance between any two different design points  $a_i, a_j \in A$ . The  $p$ -norm is typically used to calculate  $d$ :

$$d(a_i, a_j) = \|a_i - a_j\|_p = \left( \sum_{k=1}^m |a_{ik} - a_{jk}|^p \right)^{1/p} \quad (2)$$

where, for example, if  $p = 2$  then  $d$  is the Euclidean distance.

A *nested* factor is one that only exists for a specific level of an additional qualitative factor called a *branching* factor. Hung et al. (2009) proposed a *Branching Latin Hypercube Design* (BLHD) to handle an experiment with branching and nested factors as well as shared factors across all levels of the branching factors. For example, consider evaluating the performance

of a machine cutting tool that can be produced in two distinct styles. The second machine cutting tool style adds a handle attachment to the tool that has an angle and length measure, and this handle attachment is absent from the first style of the cutting tool. In this scenario, the style of the cutting tool,  $s$ , is a categorical branching factor, and the angle,  $a_1$ , and length,  $a_2$ , of the handle attachment are continuous nested factors when  $s = 2$ . In addition, the length,  $b_1$ , and width,  $b_2$ , of the cutting tool are shared factors that exist for both styles.

In order to progress to the computer experiment scenario this paper is concerned with, we introduce two additional layers of complexity:

1. The branching factors are continuous rather than categorical. Instead of possessing an explicitly known range for a continuous nested factor that exists within a specific level of a categorical branching factor, we have a nested factor whose range depends on the value selected for a continuous branching factor. The function  $f : [\min x, \max x] \rightarrow [\min y, \max y]$  that maps the range for the branching factor,  $x$ , to the range for the nested factor,  $y$ , has no closed form solution.
2. There are more than two layers of nesting. While a continuous factor is nested within a continuous branching factor, the nested factor is itself a branching factor for an additional continuous nested factor. The number of nesting layers is arbitrary, but we will make the assumption that only one factor is nested for every layer. Therefore, if we have a system of  $k$  continuous factors where each factor is nested within one other factor, then there are  $k - 1$  functions  $f_1, f_2, \dots, f_{k-1}$  that map each branching factor to its subsequent nested factor.

Figure 1 provides an illustration of our multi-layered continuous nested factors relationship. In Figure 1, we use the notation described previously along with  $u_i$  for  $i = 1, 2, \dots, k$  being the selected value for the  $i$ -th factor in the  $k$ -branching/nested factor system. For the remainder of this paper, we will use  $(u_1, u_2, \dots, u_k) = \vec{u}$  as notation for a single point in the space,  $U$ , of all possible values for the multi-layered continuous nested factors.

$$\begin{aligned}
u_1 &\in [\min u_1, \max u_1] \\
&\quad \Downarrow f_1 \\
u_2 &\in [\min u_2, \max u_2] \\
&\quad \Downarrow f_2 \\
u_3 &\in [\min u_3, \max u_3] \\
&\quad \vdots \\
u_{k-1} &\in [\min u_{k-1}, \max u_{k-1}] \\
&\quad \Downarrow f_{k-1} \\
u_k &\in [\min u_k, \max u_k]
\end{aligned}$$

Figure 1: A system of multi-layered continuous nested factors where each  $f_1, f_2, \dots, f_{k-1}$  has no closed form solution.

The remainder of this article proceeds as follows. In the next section, we discuss the aerospace simulation study that served as the impetus for this work. Next, we formally introduce the approach taken to construct a multi-NFD and subsequently illustrate the design and analysis for our specific application. The article concludes with a summary and brief discussion.

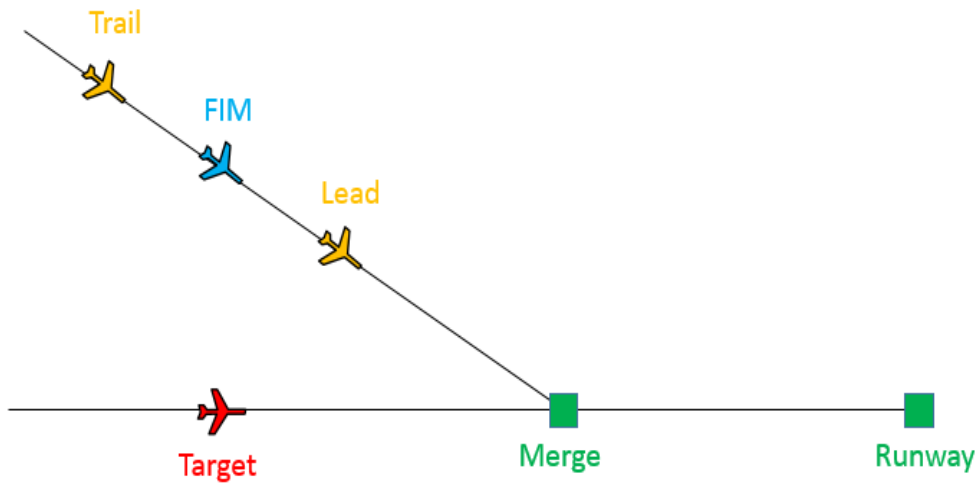
## Motivation

In this section, we discuss a specific aerospace simulation study that presented a system of multi-layered continuous nested factors. Within the National Aeronautics and Space Administration (NASA)'s Airspace Operations and Safety Program, the Air Traffic Management Technology Demonstration # 1 (ATD-1) is a proposed solution in response to current aircraft arrival procedures leading to high demand on Air Traffic Control, increased delays, and reduced capacity airports during peak periods of commercial air travel. ATD-1 is developing technologies to resolve these issues through an integrated system of airborne and ground-based technologies and decision support tools. Its purpose is to demonstrate the operational capabilities of three new integrated research technologies. One of these is *Flight Deck-based*

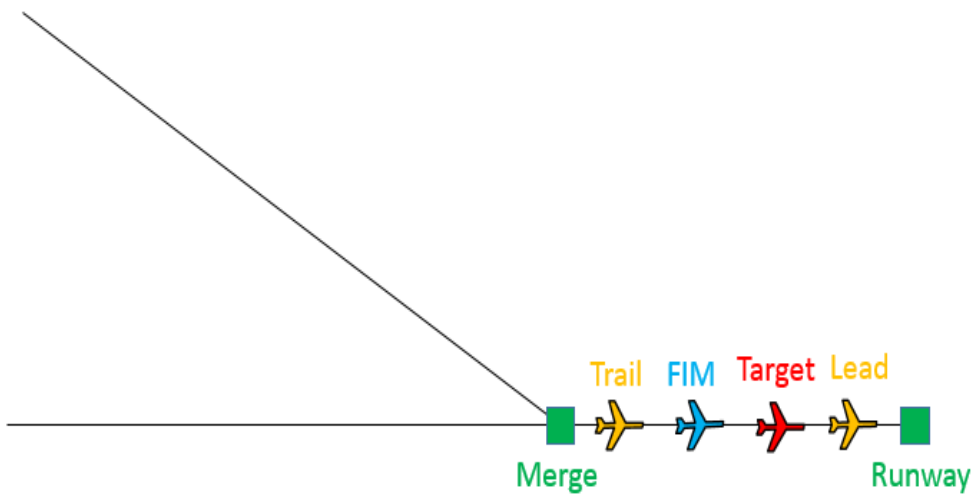
*Interval Management* (FIM): a system of aircraft avionics and flight crew procedures for achieving or maintaining aircraft spacing requirements. FIM is a speed control technology that allows an appropriately equipped aircraft, which we call the FIM aircraft, to achieve a desired spacing goal behind a target aircraft at a specified location usually within five nautical miles (nmi) of the arrival airport runway (Baxley et al., 2013).

A FIM operated arrival consists of the following steps: the FIM and target aircraft schedules are computed, the FIM aircraft is given a clearance by Air Traffic Control to begin FIM operations, the FIM aircraft flies speeds produced by the FIM speed-control algorithm, and the FIM aircraft achieves a spacing goal behind the target aircraft at a specified location. The case we are specifically interested in is where the FIM and target aircraft are flying separate arrival routes that merge together just prior to the runway. In this scenario, the FIM aircraft achieves a desired spacing goal behind the target aircraft at the point where the two routes merge, which we call the merge point. Prior to the merge point, FIM is flying in between two non-FIM aircraft with speeds controlled by air traffic controllers: a lead and trail aircraft. However, after the merge location, the FIM aircraft begins to fly immediately behind the target aircraft instead of the lead aircraft. Figure 2 (a) and (b) illustrates this particular aircraft configuration before and after the merge point, respectively.

Previous human-in-the-loop (HITL) experiments involving pilots and air traffic controllers indicated that spacing violations between aircraft can occur more frequently with FIM operations in this scenario, specifically between the FIM and trail aircraft prior to the merge point. These HITL experiments also showed that certain aircraft delay assignments and, more specifically, the difference in delay assignments between the FIM aircraft and non-FIM aircraft can lead to an increased occurrence of spacing violations (Robinson, 2014). In this context, ‘delay’ refers to the amount of time in seconds that an aircraft is estimated to arrive ahead of schedule. Aircraft absorb this delay by decreasing their speed throughout the arrival route in order to meet the schedule. A particular aircraft’s Scheduled Times of Arrival are computed based on those of the preceding aircraft in order to adhere to spacing



(a)



(b)

Figure 2: The simulated airspace configuration prior to the merge point (a) and after (b).

requirements.

A computer simulation environment was developed to simulate the previously described FIM arrival scenario, and a computer experiment was needed to investigate to what extent differences in delay assignments influence the spacing between aircraft during an arrival. This complex aircraft scheduling and delay assignment process leads to five factors: the delay initially assigned to the lead aircraft and the starting location of all four aircraft referred to as the initial distance-to-go ( $DTG_i$ ) to the runway in nmi. The order in which we can select the values for these five factors is: 1) Lead aircraft  $DTG_i$ , 2) Lead aircraft delay, 3) Target aircraft  $DTG_i$ , 4) FIM aircraft  $DTG_i$ , 5) Trail aircraft  $DTG_i$ .

Thus, these five continuous factors have a multi-layered nested relationship. The particular value selected for each factor directly affects the range of possible values for the subsequent factor by the nature of the aircraft scheduling procedure. However, the specific relationships have no closed form solution. In this particular system of multi-layered continuous nested factors, the lead aircraft  $DTG_i$  is the root factor with an unchanging range.

## Challenges

In addition to the complexities listed above (i.e., continuous branching factors, multiple layers of continuous nested factors, and no closed form solution for the relationship between the layers of nested factors), our specific application involving multi-layered continuous nested factors also faced two additional challenges. First, the actual factors of interest from a modeling perspective (in terms of their effect on the response) were the differences in delay assignments between the FIM aircraft and the three other aircraft. These three factors, which we call the *delay differentials*, directly depend on the values selected for the five continuous nested factors in the aircraft scheduling process. However, this specific relationship also has no closed form solution. Figure 3 illustrates this relationship with  $S$  being the unknown function that maps points in the continuous nested factor space of aircraft scheduling



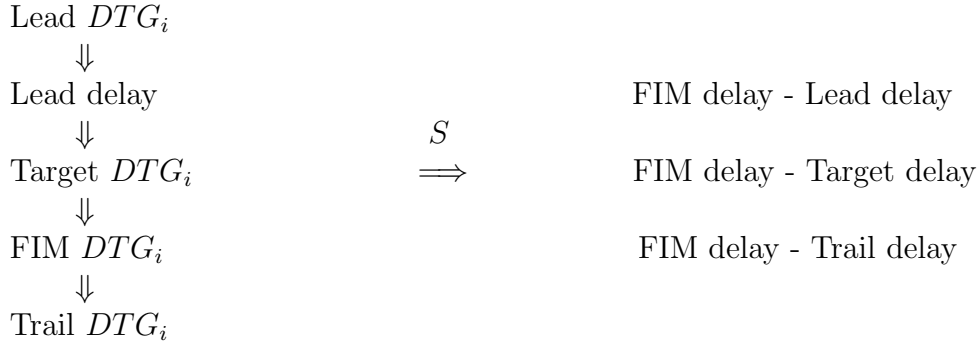


Figure 3: The specific system of five multi-layered continuous nested factors mapped by an aircraft scheduling procedure to three aircraft delay differentials.

factors to the delay differentials. Second, winds were an additional factor incorporated into the computer experiment. Wind speeds and directions are used in the calculation of aircraft speeds and Scheduled Times of Arrival, which impacts the branching/nested relationships within the aircraft scheduling factors. To combat this, we sought a space-filling method that performed well for most winds.

## Space-Filling Approach

In this section, we propose a simple approach for constructing a space-filling design with a system of  $k$  multi-layered continuous nested factors where the factor relationships between nesting layers have no closed form solution. In the next section, we discuss how we adjusted this approach to handle our specific aerospace application and challenges.

The proposed space-filling approach hinges on selecting the value,  $u_1$ , for the root branching factor in the branching/nested factor relationship. For choosing the number of points to be sampled by the space-filling method, we reference Loepkky et al. (2009) who propose an informal rule for the number of runs to be conducted in a deterministic computer experiment: ten times the dimension of the factor space,  $k$ . An approach for selecting  $10k$  design points  $(u_1, u_2, \dots, u_k) = \vec{u} \in U$  is as follows:

1. Divide the range of the root branching factor,  $[\min u_1, \max u_1]$ , into  $J$  closed intervals of equal length where  $1 \leq J \leq 10k$  and  $10k$ , the number of design points, is a multiple of  $J$ . In order to cover the entire range of the root factor and have the intervals be of the same length, the intervals are disjoint partitions with the exception of their endpoints. Let  $i$  specify one of the  $J$  equally sized interval partitions of  $[\min u_1, \max u_1]$ , and set  $i = 1$  initially.
2. Select  $u_1$  randomly from the  $i$ th interval. That is, the value for the root branching factor is drawn from the uniform distribution:

$$u_1 \sim \text{unif}\left[\min u_1 + \frac{i-1}{J}(\max u_1 - \min u_1), \min u_1 + \frac{i}{J}(\max u_1 - \min u_1)\right].$$

3. Select the values of all remaining branching/nested factors  $u_2, u_3, \dots, u_k$  randomly from their resulting ranges in sequential order:

$$u_2 \sim \text{unif}[\min u_2, \max u_2]$$

$$u_3 \sim \text{unif}[\min u_3, \max u_3]$$

$$\vdots$$

$$u_k \sim \text{unif}[\min u_k, \max u_k].$$

4. Repeat Steps 2-3 until there are  $\frac{10k}{J}$  sample design points originating from the  $i$ -th interval partition of the root branching factor range.
5. Repeat Steps 2-4 each time incrementing  $i$  by one until  $\frac{10k}{J}$  design points originate from sampling  $u_1$  randomly from each of the  $J$  partitions of its range for a total of  $10k$  design points.

This base design method serves as an initial approach for a multi-NFD, which can be adjusted for a specific computer experiment application by choosing the value of  $J$ . The

initial base design method can also be adjusted by dividing the range of  $u_1$  into  $J$  intervals of unequal length, and a different number of points can originate from sampling  $u_1$  from within each of its range's  $J$  intervals. These adjustments are intended to handle the complex interdependencies between the factors within the specific computer experiment with the assumption that the root factor has the most influence over the resulting ranges of the remaining branching/nested factors. Adjusting the base design method is intended to sample rare, yet possible, design points and to uniformly fill a possibly irregularly shaped design region,  $U \subset \mathbb{R}^k$ . If time permits, we recommend testing a number of base design methods (with different adjustments) to determine which method works best for the specific application. In order to compare base design method alternatives and select the 'best' space-filling design for a specific application, the maximin criterion can be used:

1. For a given base design approach, generate many (possibly thousands) of designs of  $10k$  points.
2. For each design, compute the maximin criterion (equation 1).
3. Average this criterion over all designs.
4. Repeat Steps 1-3 for each adjustment of the base design method, and select the one that maximizes the minimum distance between every pair of design points.

Once the base design approach method has been selected, use the maximin criterion again to select the 'best' design of  $10k$  points produced by the chosen method, and use these design points to run the computer experiment.

## Application to Aerospace Simulation Study

We now describe how the general multi-NFD procedure was adjusted for our specific aerospace application and then briefly describe the analysis conducted.

## Design

As mentioned previously, our computer experiment consisted of five multi-layered nested factors (see Figure 3), which were mapped by the scheduling procedure, function  $S$ , to the FIM aircraft's three-dimensional delay differential space. In order to explore the space of possible values for the delay differentials and better understand how  $S$  behaves, a fine grid of combinations of values for the five branching/nested factors were selected and the resulting delay differentials were plotted. Eight different wind conditions were used in this simulation study, and, for brevity, we show the delay differential space for wind #1. Figure 4 illustrates the delay differential space for wind # 1 in three dimensions (a) and projected onto each two-dimensional plane (b)-(d).

Figure 4 shows that the delay differential space is more dense towards the middle of its region and becomes less dense towards its extreme values. Studying the effect of extreme delay differences on aircraft spacing was of interest to the researchers. We emphasize again that, for this particular application, we are looking to construct a space filling design in the delay differential space rather than the nested factor space. Since the delay differential space is of interest from a modeling perspective, its dimension was used in the calculation of the number of design points to be used in the computer experiment:  $10 \times 3 = 30$  design points.

A number of possible multi-NFD methods were investigated using the adjustments to the base design method discussed in the previous section, including those with unequal splits of the root factor range and an unequal number of design points sampled from each interval. Through exploration, we found that the minimum distance between every pair of design points was large when the space-filling method consisted of dividing the root factor range into equally sized intervals with an equal number of points sampled from each interval. Thus, the multi-NFD methods compared for this application only differed in the number of equally sized intervals of the root factor's range:

- *No Split* ( $J=1$ ): In this design method, all points are randomly selected from the full range of values for the root factor, and the values for the remaining factors are selected randomly from their full resulting ranges.
- *Split Half* ( $J=2$ ): The range of the root branching factor is divided into two equally-sized intervals, and one half of the design points start with  $u_1$  randomly selected from each interval. The values for the remaining factors  $u_2, u_3, u_4$ , and  $u_5$  are chosen randomly within their full resulting ranges.
- *Split Third* ( $J=3$ ): The range of the root branching factor is divided into three equally-sized intervals, and one third of the design points start with  $u_1$  randomly selected from each interval. The values for the remaining factors  $u_2, u_3, u_4$ , and  $u_5$  are chosen randomly within their full resulting ranges.
- *Split Fifth* ( $J=5$ ): The range of the root branching factor is divided into five equally-sized intervals, and one fifth of the design points start with  $u_1$  randomly selected from each interval. The values for the remaining factors  $u_2, u_3, u_4$ , and  $u_5$  are chosen randomly within their full resulting ranges.

We compared the four selected space-filling methods within each wind condition using the maximin criterion. For a given wind condition, each of the four techniques was used to sample 1000 designs of 30 points. For each design, the minimum distance between every pair of points was computed, and this value was averaged across all 1000 designs. Table 1 presents the average minimum distance between every pair of design points for each of the four selected techniques for all eight wind conditions.

Although the differences are not substantial, Table 1 shows that in terms of the maximin criterion, the *Split Fifth* space-filling method performed the best or second-best for six out of eight wind conditions. After selecting *Split Fifth*, the maximin criterion was used to select the ‘best’ design of 30 points for each of the eight wind conditions. Figure 4 shows the 30 delay differential design points selected by the *Split Fifth* method for wind # 1. Plots (b)-(d)

Table 1: The average  $\min d(\cdot, \cdot)$  over 1000 iterations of each space-filling technique for all wind sets. \* Largest average  $\min d(\cdot, \cdot)$  for a given wind condition.

<b>Wind \ Technique</b>	<i>No Split</i>	<i>Split Half</i>	<i>Split Third</i>	<i>Split Fifth</i>
Wind # 1	8.3417*	8.3057	8.1707	8.2772
Wind # 2	7.7411	7.4806	7.6174	7.7561*
Wind # 3	7.9675	7.9225	7.9879	7.990*
Wind # 4	9.2125*	9.1343	8.8717	9.1722
Wind # 5	8.6486	8.7715	8.5702	8.8082*
Wind # 6	7.0062	7.2995*	7.2465	7.1929
Wind # 7	8.2377	8.1586	8.1751	8.2576*
Wind # 8	7.6751	7.6574	7.7306*	7.7088

in Figure 4 show the 30 design points projected onto each two-dimensional delay differential plane. The points are plotted on top of the full delay differential space projections for wind # 1. The design points appear to fill the space reasonably well, and we were successful in sampling some points from the extreme design regions where particular combinations of delay differential values are less common.

## Analysis

This section describes results from a Gaussian process model (Brockwell and Davis, 1991) fit from the output data of our aerospace computer simulation experiment. This analysis technique is a very common method used for space-filling designs with a deterministic computer experiment. A Gaussian process model creates a predicted response surface over the entire experimental design region using interpolation. This model uses the assumption that any finite collection of responses in the design region is distributed as a multivariate Gaussian to predict the response at untested design points. In our case, the response was the *Closest Point of Approach (CPA)*, which is a measure of the minimum distance between the FIM aircraft and the lead and trail aircraft over the course of an arrival (throughout one exper-

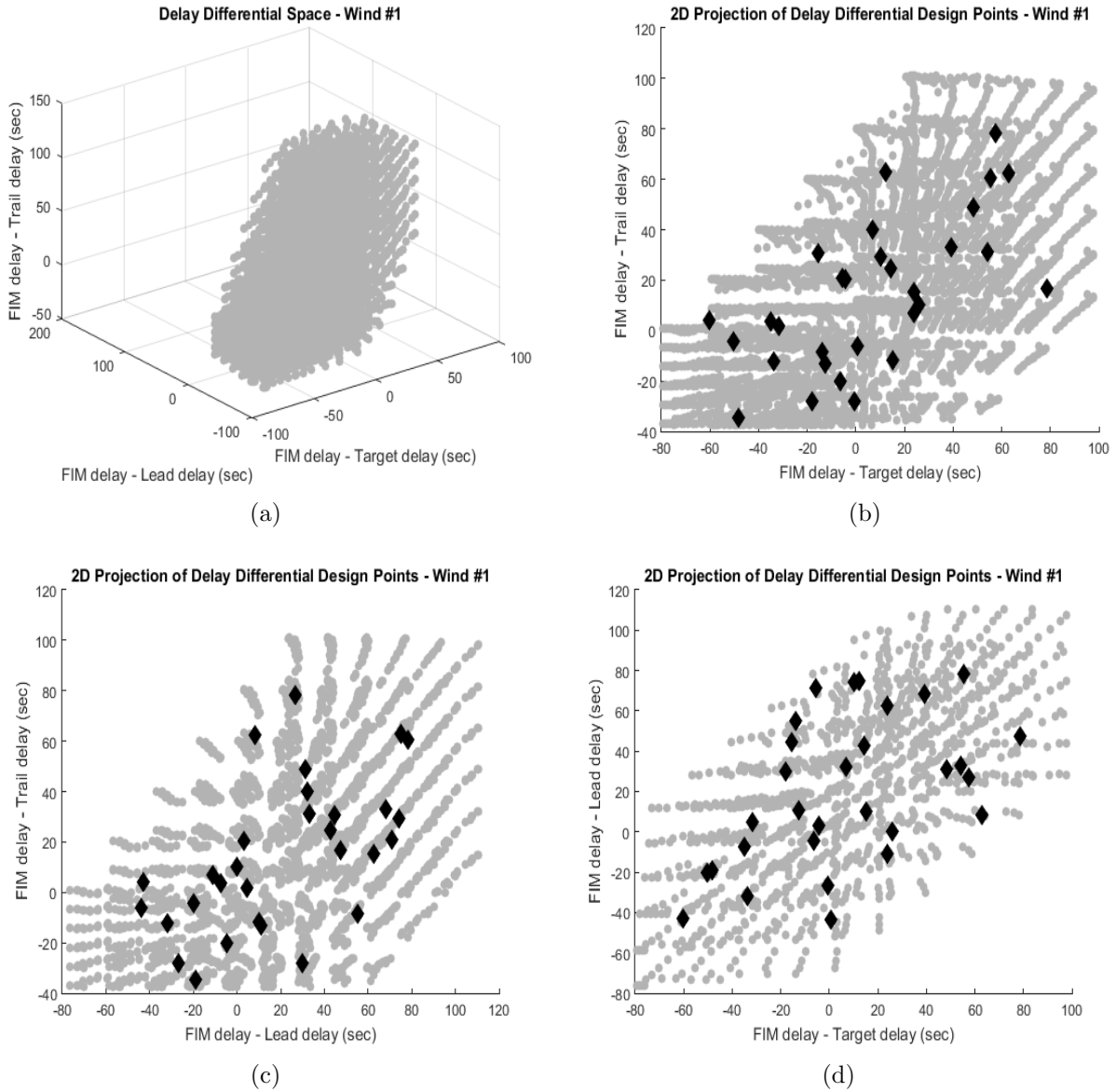


Figure 4: The full 3-D delay differential space for wind # 1 (a) along with the 30 design points sampled from the *Split Fifth* space-filling method projected onto each two-dimensional delay differential plane (b)-(d).

imental run). The *CPA* is a continuous response with a negative value corresponding to a spacing violation, i.e., the aircraft do not meet the minimum required separation standard. See the Appendix for a mathematical derivation of the particular Gaussian process model used for our analysis.

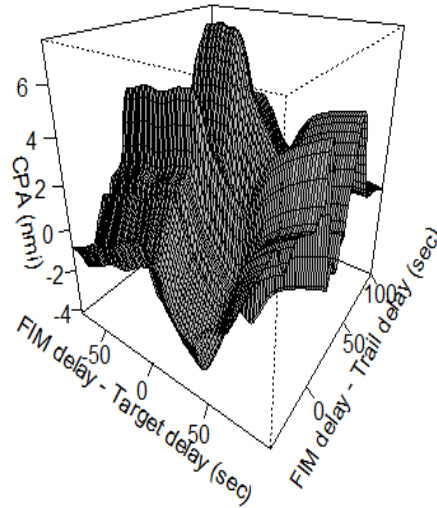
Figures 5-7 display the *CPA* response surface predicted by the Gaussian process model against each pair of delay differentials at two different angles for visualization. Only results for wind # 1 are shown since predicted surfaces were similar for all wind conditions. The *CPA* response surface shows how differences in delay assignments affect distance spacings between aircraft during an arrival. When interpreting the surface, keep in mind that small (negative) *CPA* values are undesired as they correspond with spacing violations. Upon inspection of the surface, we see that very positive and negative *CPA* values can be achieved by many combinations of aircraft delay differentials. However, a close look at Figure 5 reveals that a minimum *CPA* is achieved at a FIM and target aircraft delay differential of around 50 seconds across all values of the FIM and trail aircraft delay differential. That is, a spacing violation is likely to occur when the FIM aircraft is assigned approximately 50 more seconds of delay than the target aircraft. Also, from Figure 6, we see that overall, *CPA* decreases as the FIM and lead aircraft delay differential increases. That is, spacing violations are more likely to occur as the FIM aircraft is assigned more delay than the lead aircraft. This is valuable information for future research on FIM operations during similar arrival scenarios.

## Discussion

In this paper, we describe the problem of designing a computer experiment when there are multiple layers of continuous nested factors with no closed form solutions for their relationships. We provide a motivating example from an aerospace computer simulation experiment which had some additional challenges. We propose a simple approach for constructing a

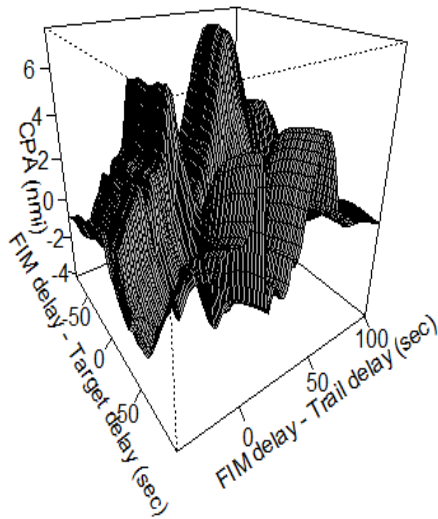


**CPA Response Surface vs FIM/Target and FIM/Trail Delay Differentials - Angle #1**



(a)

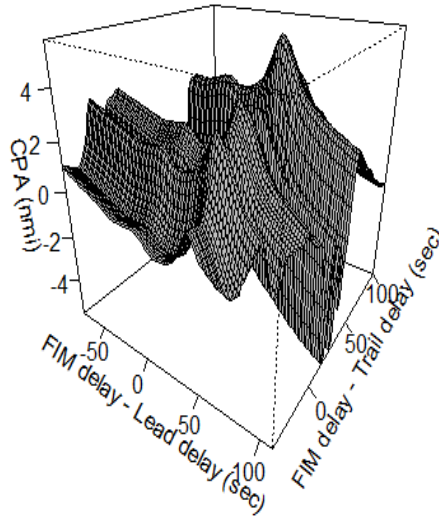
**CPA Response Surface vs FIM/Target and FIM/Trail Delay Differentials - Angle #2**



(b)

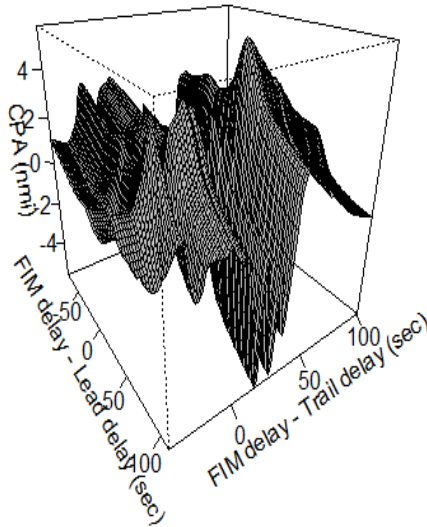
Figure 5: The predicted *CPA* response surface versus the FIM and target aircraft delay differential and the FIM and trail aircraft delay differential at two different angles.

**CPA Response Surface vs FIM/Lead and FIM/Trail Delay Differentials - Angle #1**



(a)

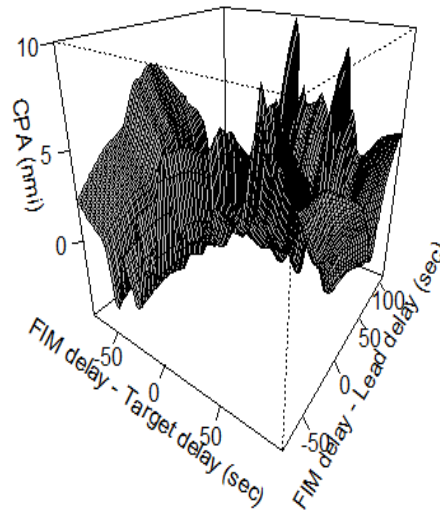
**CPA Response Surface vs FIM/Lead and FIM/Trail Delay Differentials - Angle #2**



(b)

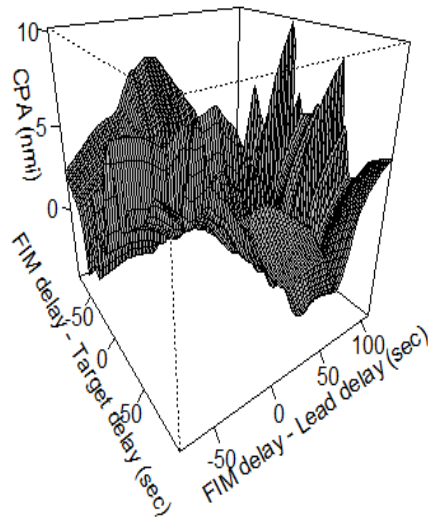
Figure 6: The predicted *CPA* response surface versus the FIM and lead aircraft delay differential and the FIM and trail aircraft delay differential at two different angles.

**CPA Response Surface vs FIM/Target and FIM/Lead Delay Differentials - Angle #1**



(a)

**CPA Response Surface vs FIM/Target and FIM/Lead Delay Differentials - Angle #2**



(b)

Figure 7: The predicted *CPA* response surface versus the FIM and target aircraft delay differential and the FIM and lead aircraft delay differential at two different angles.

space-filling design with a system of multi-layered continuous nested factors, a multi-NFD, that can be adapted for specific computer experiments. The maximin criterion can be used to select the best multi-NFD method adaptation as well as the best design produced by the chosen method. We illustrate the proposed approach with our specific application and present results from a Gaussian process model fit to predict a response surface. Results from this computer experiment inform researchers developing procedures for when FIM operations should be conducted to minimize the risk of spacing violations between aircraft. Our approach can be used for future computer simulation experiments involving unique factor relationships of this type, which NASA has seen in previous studies and predicts to see more of in the future.

For future research, we recommend investigating more structured design methodologies for multi-layered continuous nested factors. For example, the value for the root branching factor could be set and the design approach begin at the subsequent factor effectively making it the new root branching factor. Other criteria for space-filling designs can also be investigated such as the minimum potential or maximum entropy criterion. Additional factor relationships similar to multi-layered continuous nested factors are also worth further research; these can include more than one continuous factor nested within a particular continuous branching factor as well as the simpler case of only one layer of a branching and nested factor.

## Appendix

Let  $y$  be the function that maps the delay differential input variable space,  $D$ , to the  $CPA$  output response space,  $C$ , within the computer experiment as such:

$$y : D \rightarrow C, \quad (3)$$

$$D \subset \mathbb{R}^3,$$

$$C \subset \mathbb{R}.$$

For the 30 delay differential design points  $d_1, d_2, \dots, d_{30} \in D$  selected by the space-filling method, the *Split Fifth* multi-NFD method, we know  $y(d_1), y(d_2), \dots, y(d_{30})$  as the *CPA* responses recorded by the computer simulation. But what about  $y(d^*)$  for some untested delay differential  $d^* \in D$ ? We estimate this value with a *Gaussian process model*, or GP model. The following formulation for a GP model follows MacDonald et al. (2015) and Do (2007).

A *Gaussian process* is a stochastic process where any finite collection of random variables is distributed as a multivariate Gaussian. That is, a finite collection of random variables  $\{h(d) : d \in D\}$  is distributed as a Gaussian process with mean function  $m$  and covariance function  $k$  if for any finite set of elements from the index set  $d_1, d_2, \dots, d_n \in D$ , the set of random variables  $h(d_1), h(d_2), \dots, h(d_n)$  is distributed as

$$\begin{bmatrix} h(d_1) \\ \vdots \\ h(d_n) \end{bmatrix} \sim N \left( \begin{bmatrix} m(d_1) \\ \vdots \\ m(d_n) \end{bmatrix}, \begin{bmatrix} k(d_1, d_1) & \cdots & k(d_1, d_n) \\ \vdots & \ddots & \vdots \\ k(d_n, d_1) & \cdots & k(d_n, d_n) \end{bmatrix} \right),$$

which we also write as  $h(\cdot) \sim GP(m(\cdot), k(\cdot, \cdot))$ .

In order to connect this GP model formulation with our specific problem, we can model the *CPA* response as a realization of the following stochastic process:

$$y(d_i) = \mu + h(d_i) \tag{4}$$

$$i = 1, 2, \dots, 30$$

where  $d_i$  is one of the 30 delay differential design points,  $h(\cdot) \sim GP(m(\cdot), k(\cdot, \cdot))$  as previously defined, and  $\mu$  is the true process mean. This is the same function that maps the delay differential input variable space,  $D$ , to the  $CPA$  output response variable space,  $C$ , in Equation (4). Even though the computer simulation and the input to output relationships are deterministic, we have formulated the recorded responses as realizations of the stochastic process that is the physical process being emulated, which contains some randomness and uncertainty. This gives us a probabilistic means of predicting the response from a deterministic computer experiment.

We will make the following common assumptions about the mean and covariance functions of the GP process  $h(\cdot)$ :

$$m(h(d_i)) = E(h(d_i)) = 0 \tag{5}$$

$$k(h(d_i), h(d_i)) = Var(h(d_i)) = \sigma^2 \tag{6}$$

$$k(h(d_i), h(d_j)) = Cov(h(d_i), h(d_j)) = \sigma^2 R_{ij} \tag{7}$$

to allow for ease in parameter estimation. Due to how  $y(\cdot)$  is constructed from the  $h(\cdot)$  GP process by adding a constant mean  $\mu$  from Equation (5), the assumptions in Equations (6)-(8) imply that  $y(\cdot)$  is also a GP process:

$$y(\cdot) \sim GP(\mu, \sigma^2 \mathbf{R})$$

where  $\mathbf{R}$  is the covariance matrix with entries  $R_{ij}$ . If we make a 30 x 30 data matrix  $\mathbf{D}$

from our 30 delay differential points  $d_1, d_2, \dots, d_{30} \in D$ , then the distribution of  $y(\cdot)$  allows  $\mathbf{Y} = y(\mathbf{D}) = (y(d_1), y(d_2), \dots, y(d_{30}))^T$  as the vector of 30 *CPA* responses to be distributed as a multivariate Gaussian:

$$\mathbf{Y} = y(\mathbf{D}) \sim N(\mathbf{1}\mu, \sigma^2\mathbf{R})$$

where  $\mathbf{1}$  is a 30 x 1 vector of one's. We choose the most commonly used Gaussian correlation function for  $\mathbf{R}$ . Also known as the *squared-exponential correlation* function, it is given by

$$R_{ij} = \prod_{k=1}^3 \exp\{-\theta_k |d_{ik} - d_{jk}|^2\} \quad (8)$$

where  $\theta = (\theta_1, \theta_2, \theta_3)$  is a vector of parameters that require estimation. This common correlation structure is popular due to its desirable properties. These include local smoothness; that is, nearby  $d_i$ 's will have a correlation  $\approx 1$  while  $d_i$ 's that are far apart will have a correlation  $\approx 0$ . This means that if an untested delay differential point is very close to a tested point, then the predicted response at the untested point will be very close to the recorded response at the tested point.

Using the fact that  $\mathbf{Y} \sim N(\mathbf{1}\mu, \sigma^2\mathbf{R})$ , we can derive the Maximum Likelihood Estimators (MLE)

$$\hat{\mu} = (\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1})^{-1} (\mathbf{1}^T \mathbf{R}^{-1} \mathbf{Y}) \quad (9)$$

$$\hat{\sigma}^2 = \frac{1}{30} \cdot (\mathbf{Y} - \mathbf{1}\hat{\mu})^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{1}\hat{\mu}) \quad (10)$$

for the  $\mu$  and  $\sigma^2$  parameters. The MLE estimates in Equations (10) and (11) are used to obtain the following negative log-likelihood function

$$-2\log(L_\theta) \propto \log(|\mathbf{R}|) + 30 \cdot \log[(\mathbf{Y} - \mathbf{1}\hat{\mu})^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{1}\hat{\mu})] \quad (11)$$

for estimating the  $\theta$  parameters where  $|\mathbf{R}|$  is the determinant of  $\mathbf{R}$ . Using the MLE technique along with the parameter estimates and negative log-likelihood function in Equations (10)-(12), we obtain the best linear unbiased estimator of the *CPA* response at an untested  $d^* \in D$  as

$$\hat{y}(v^*) = \hat{\mu} + \mathbf{r}^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{1}\hat{\mu}) = \left[ \frac{(1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{1})}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}} \mathbf{1}^T + \mathbf{r}^T \right] \mathbf{R}^{-1} \mathbf{Y} = \mathbf{Z}^T \mathbf{Y} \quad (12)$$

where we make the following substitution:

$$\mathbf{Z}^T = \left[ \frac{(1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{1})}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}} \mathbf{1}^T + \mathbf{r}^T \right] \mathbf{R}^{-1} \quad (13)$$

in which  $\mathbf{r} = (r_1(d^*), r_2(d^*), \dots, r_{30}(d^*))$  with  $r_i(d^*) = \text{corr}(h(d^*), h(d_i))$  computed by plugging  $d^*$  and  $d_i$  into Equation (9) for  $R_{ij}$ . The predicted response at an untested design point in Equation (13) also has a Mean Square Error of

$$\begin{aligned} MSE(\hat{y}(d^*)) &= E[(\hat{y}(d^*) - y(d^*))^2] \\ &= \hat{\sigma}^2 (1 - 2\mathbf{Z}^T \mathbf{r} + \mathbf{Z}^T \mathbf{R} \mathbf{Z}) = \hat{\sigma}^2 \left[ 1 - \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} + \frac{(1 - \mathbf{1}^T \mathbf{R}^{-1} \mathbf{r})^2}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}} \right] \end{aligned}$$

where we take advantage of Equation (14) for easier notation. This is how we estimate the *CPA* response at untested delay differential points using the recorded responses from the 30 tested points. We can now predict a coarse *CPA* response surface over a grid of points covering the entire three-dimensional delay differential space,  $D$ .



## References

- Baxley, B., Johnson, W., Swenson, H., Robinson, J., Prevot, T., Callantine, T., Scardina, J., and Greene, M. (2013). “Air Traffic Management Technology Demonstration-1 Concept of Operations (ATD-1 ConOps), Version 2.0”. *National Aeronautics and Space Administration*.
- Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods, 2nd ed.* Springer-Verlag, New York, NY.
- Dean, A., Morris, M., and Stufken, J. (2015). “Computer Experiments”. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods: Handbook of Design and Analysis of Experiments* (pp. 577-626), CRC Press.
- Do, C. (2007). “Gaussian Processes”. Retrieved from <https://see.stanford.edu/materials/aimlcs229/cs229-gp.pdf>.
- Hung, Y., Joseph, R., and Melkote, S. (2009). “Design and Analysis of Computer Experiments With Branching and Nested Factors”. *Technometrics*, 51, 4, pp. 354-365.
- Johnson, M., Moore, L., and Ylvisaker, D. (1990). “Minimax and Maximin Distance Designs”. *Journal of Statistical Planning and Interface*, 26, pp. 131-148.
- Loepky, J., Sacks, J., and Welch, W. (2009). “Choosing the Sample Size of a Computer Experiment: A Practical Guide”. *Technometrics*, 51, 4, pp. 366-376.
- MacDonald, B., Ranjan, P., and Chipman, H. (2015). “GPfit: An R Package for Fitting a Gaussian Process Model to Deterministic Simulator Outputs”. *Journal of Statistical Software*, 64, 12.
- McKay, M., Beckman, R., and Conover, W. (1979). “A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer

Code”. *Technometrics*, 21, 2, pp. 239-245.

- Montgomery, D. (2012). “Experiments with Computer Models”. *Design and Analysis of Experiments, 8th ed.* (pp. 523-530), Wiley Global Education.
- Qian, P. and Wu, J. (2009). “Sliced Space-Filling Designs”. *Biometrika*, 96, 4, pp. 945-956.
- Robinson, J. (2014). “Calculation of Flight Deck Interval Management Assigned Spacing Goals Subject to Multiple Scheduling Constraints”. *33rd Digital Avionics Systems Conference*.