Factor Effects in Numerical Simulations

JUDAH L. CLEVELAND^a AND JEFFREY A. SMITH

Army Research Laboratory, U.S. CCDC, White Sands Missile Range, New Mexico

JAMES P. COLLINS

Army Research Laboratory, U.S. CCDC, Adelphi, Maryland

(Manuscript received 24 September 2019, in final form 17 April 2020)

ABSTRACT

Numerical simulations allow users to adjust factor settings in experimental runs to understand how changes in those factors affect the output. However, it is not straightforward to analyze these outputs when multiple input factors are changed, especially simultaneously. For the atmospheric sciences, Stein and Alpert introduced a method they termed “factor separation” in order to separate the “pure contribution” of a factor from “pure interactions” of combinations of factors. Although factor separation appears to be used exclusively within the atmospheric sciences, other communities achieve a similar result by computing “main effects” via design of experiments methods. While both methods yield different estimates for the factor effects or contributions, we show that factor separation effects are identical to “simple effects” in the design of experiments literature. We demonstrate how both factor separation effects and design of experiments main effects correspond to multiple linear regression coefficients with different coding methods; thus, effect estimates produced by each method are equivalent through a variable transformation. We illustrate the application of both methods using a shallow-water simulation. This connection between factor separation and the design of experiments discipline extends factor separation to more applications by making available design of experiments methods for decreasing the computational cost and calculating effects for factors with more than two settings, both of which are limitations of factor separation.

1. Introduction

A common theme in many experiments is a desire to both identify and quantify relationships between various factors of interest and the output. Within the atmospheric science community, factor separation (FS) (Stein and Alpert 1993) is one method of conducting such an analysis; however, in communities outside the atmospheric sciences, main and interaction effects from design of experiments (DoE) methods are widely used (e.g., Wu and Hamada 2009). Within the atmospheric science community a number of studies use FS methods; however, to the best of our knowledge DoE has seen little formal practice save only recent work (Smith and Penc 2016; Smith et al. 2018, 2019), which

applied DoE methods to the study of a numerical weather prediction code.

To compute the contributions of various factors and their interactions to a predicted field, Stein and Alpert (1993) developed factor separation. In particular, they emphasize that neglecting the presence of interaction between factors can yield misleading results (Alpert and Sholokhman 2011a). Stein and Alpert (1993) demonstrated FS by using it to study the effects of surface fluxes and terrain on precipitation in a simulation. Since then, factor separation has been applied to temperature–albedo feedback of the greenhouse gases (Berger et al. 2011); the effects of topography, convection, latent, and sensible heat fluxes on Alpine lee cyclogenesis (Alpert 2011); the effects of waste heat, vapor, and pollution on cumulus convection (Reuter 2011); and many other situations (see Alpert and Sholokhman 2011a). Factor separation is such a simple and practical method for isolating the contributions and interactions of various factors that members of the atmospheric community quickly saw its usefulness and applied it to many

^a Current affiliation: Research Associateship Program, U.S. Army Research Laboratory, White Sands Missile Range, New Mexico.

Corresponding author: Jeffrey A. Smith, jeffrey.a.smith1.civ@mail.mil

different atmospheric simulations. More than 25 years later, the original FS paper continues to receive citations (e.g., [Thunis et al. 2019](#); [Waugh et al. 2019](#); [Yang et al. 2019](#)); however, FS does not appear to be used outside the atmospheric science community.

In other disciplines, an alternative method for attributing output response to different factors comes from the statistical DoE. DoE first appeared in 1935 through the work of Fisher who placed a particular emphasis on the construction of an experiment so to optimize the ability to make inferences about the results ([Fisher 1971](#)). While Fisher began by applying his methods to experiments in the agricultural sciences, others quickly applied Fisher's methods to fields as diverse as the chemical and nuclear industries, manufacturing, and the social sciences ([National Research Council 1995](#); [Kugler et al. 2012](#)).

The DoE method for quantifying the effects of factors is via the use of main and interaction effects ([Wu and Hamada 2009](#); [Montgomery 2013](#)). While the objective of computing these quantities is similar to that of FS, the resulting effect estimates are different. This led us to ask why these methods yield different estimates and whether one method should be preferred over the other. [Alpert and Sholokhman \(2011b\)](#) briefly explored the differences between FS and DoE main effects; however, they presented the two methods as coming from completely different mathematical foundations. We found, however, that DoE "simple effects," as opposed to "main effects," are calculated in the same manner as FS effects ([Collins et al. 2014](#)). From this we conclude that, at the level of Stein and Alpert's original paper, factor separation is a special case of design of experiments.

In this paper, we first define many of the terms we use, and then we introduce full factorial designs since both methods employ this experiment paradigm. Next, we give the mathematical foundation of factor separation followed by a discussion of DoE main effects. With both methods outlined, we show the results of our research: namely, that FS effects are actually equivalent to DoE simple effects. Furthermore, through a variable transformation, we show that FS effects are essentially equivalent to DoE main effects for analytic purposes. We show how the differences between simple and main effects do not alter conclusions regarding the models created using either method. Next, we demonstrate how to interpret the results of each method with a simple example using a shallow-water equations (SWE) model. Finally, we address the assertions of [Alpert and Sholokhman \(2011b\)](#) with regard to the benefits of FS effects over main effects, and then propose how DoE can reduce the limitations of FS when computational cost and factors with more than two levels are concerns.

2. Definitions

Here we define some common DoE and FS terms.

a. Basic definitions

- Factor—A variable in an experiment.
- Level—A particular value or setting of a factor.
- Run—An experiment with the factors each set to a specific level.
- Replication—Multiple runs with the factor levels unchanged. Note: in a well designed and constructed deterministic computer simulation, replication is trivial because each response is identical.

b. Design space

- Experimental design—The planned set of runs in an experiment.
- Full factorial design—An experiment where every combination of factor levels is present.
- Balanced design—An experiment where every factor level occurs in the same number of runs.
- Orthogonal design—An experiment where all factor level combinations occur in the same number of runs.

c. Effects estimates

- Main effect—The difference between the average of all runs at the high level of a factor and the low level of that factor.
- Conditional main effect—The main effect of a factor at a given level of another factor.
- Simple effect—The conditional main effect of a factor given all other factors are held at their low levels.
- Interaction effect—The joint effect of two or more factors.
- Pure contribution—The result of factor separation that is the fraction of the response induced by a particular factor.
- Pure interaction (or synergy)—The result of factor separation that is the fraction of the response induced by a combination of factors.
- Base case (zero state)—The run with all factors off (or low).
- Control run—The run with all factors on (or high).

d. Coding

- Coding—Assigning numbers to categorical variables in order to incorporate them into a statistical model.
- Dummy coding—The method of coding that uses 0 for off (or low) and 1 for on (or high).
- Effects coding—The method of coding that uses -1 for off (or low) and 1 for on (or high).

3. Full factorial designs

Both classical DoE and FS use a *full factorial design*, though [Stein and Alpert \(1993\)](#) do not use the term in their original paper. A full factorial design has a run for every combination of *factors* and *levels*. When all k factors have two levels, typically present and absent or high and low, the design is termed a 2^k factorial because there are 2^k combinations of factors. A full 2^k factorial design is *balanced* and *orthogonal* ([Wu and Hamada 2009](#); [Montgomery 2013](#)). [Table 1](#) illustrates the full 2^k factorial design for $k = 3$ factors.

In the discussion that follows, we focus exclusively on the case when all factors have two levels. Mixed designs, where some study factors have more than two levels, are discussed by [Wu and Hamada \(2009\)](#) and [Montgomery \(2013\)](#); however, they are not a subject of this paper.

4. Stein and Alpert factor separation

[Stein and Alpert \(1993\)](#) introduced factor separation, which calculates *pure contributions* and *pure interactions*. They set up the mathematical basis as follows:

The field f depends on n factors ψ_i ($i = 1, 2, \dots, n$), where

$$\psi_i(c_i) = c_i \psi_i, \quad 0 \leq c_i \leq 1. \quad (1)$$

When a factor ψ_i is absent $c_i = 0$, and when present $c_i = 1$. The field f is a continuous function of c_i ($i = 1, 2, \dots, n$):

$$f = f(c_1, c_2, \dots, c_n). \quad (2)$$

They then decompose f through a Taylor series expansion:

$$f(c_1, c_2, \dots, c_n) = \hat{f}_0 + \sum_{i=1}^n \hat{f}_i(c_i) + \sum_{i,j=1,2}^{n-1,n} \hat{f}_{ij}(c_i, c_j) + \dots + \hat{f}_{123\dots n}(c_1, c_2, \dots, c_n), \quad (3)$$

where each $\hat{f}_{ijk\dots}(c_i, c_j, c_k, \dots)$ is identically zero if any of the variables c_i ($i = 1, 2, \dots, n$) are zero. Additionally, we can drop the ones by setting $\hat{f}_{ij\dots} = \hat{f}_{ij\dots}(1, 1, \dots)$.

The quantity $\hat{f}_0$ is known as the base case and is the value when none of the chosen factors are present. Thus, it is the portion contributed by other influences independent of the chosen factors. The quantity $\hat{f}_i$ is the pure contribution of factor i , $\hat{f}_k$ the pure contribution of factor k , and $\hat{f}_{ik}$ the pure contribution due to the pure interaction, or synergy, of factors i and k , and similarly for more factors.

For example, for the field with three factors i, j , and k where i and k are present but j is absent, the result consists of the base case, the pure contribution of i , the pure contribution of k , and the synergy of i and k :

TABLE 1. The 2^k factorial design for $k = 3$ factors at two levels. Minus indicates the factor is absent, and plus indicates it is present.

Run number	Factor A	Factor B	Factor C	Simulation result
1	—	—	—	$\hat{f}_0$
2	+	—	—	$\hat{f}_A$
3	—	+	—	$\hat{f}_B$
4	+	+	—	$\hat{f}_{AB}$
5	—	—	+	$\hat{f}_C$
6	+	—	+	$\hat{f}_{AC}$
7	—	+	+	$\hat{f}_{BC}$
8	+	+	+	$\hat{f}_{ABC}$

$$\begin{aligned} f(c_i, c_j, c_k) &= f(1, 0, 1) = f_{ik} = \hat{f}_0 + \hat{f}_i(1) + \hat{f}_j(0) + \hat{f}_k(1) \\ &\quad + \hat{f}_{ij}(1, 0) + \hat{f}_{ik}(1, 1) + \hat{f}_{jk}(0, 1) + \hat{f}_{ijk}(1, 0, 1) \\ &= \hat{f}_0 + \hat{f}_i + \hat{f}_k + \hat{f}_{ik}. \end{aligned} \quad (4)$$

A full factorial for n factors yields the following results:

$$f_0 \equiv f(0, 0, \dots, 0) = \hat{f}_0, \quad (5)$$

$$f_i = \hat{f}_0 + \hat{f}_i, \quad (6)$$

$$f_{ij} = \hat{f}_0 + \hat{f}_i + \hat{f}_j + \hat{f}_{ij}, \quad (7)$$

$$f_{ijk} = \hat{f}_0 + \hat{f}_i + \hat{f}_j + \hat{f}_k + \hat{f}_{ij} + \hat{f}_{jk} + \hat{f}_{ik} + \hat{f}_{ijk}, \quad (8)$$

$$f_{123\dots n} = \hat{f}_0 + \sum_{i=1}^n \hat{f}_i + \sum_{i=1,2}^{n-1,n} \hat{f}_{ij} + \sum_{i,j,k=1,2,3}^{n-2,n-1,n} \hat{f}_{ijk} + \dots + \hat{f}_{123\dots n}. \quad (9)$$

Notice there are 2^n equations with 2^n unknowns so each $\hat{f}_{ij\dots}$ can be solved for explicitly using recursive elimination.

For example, using the runs from the 2^3 factorial design found in [Table 1](#), the pure contributions and interactions for the three factors A, B , and C are computed by

$$\hat{f}_0 = f_0, \quad (10)$$

$$\hat{f}_A = f_A - f_0, \quad (11)$$

$$\hat{f}_B = f_B - f_0, \quad (12)$$

$$\hat{f}_C = f_C - f_0, \quad (13)$$

$$\hat{f}_{AB} = f_{AB} - f_A - f_B + f_0, \quad (14)$$

$$\hat{f}_{AC} = f_{AC} - f_A - f_C + f_0, \quad (15)$$

$$\hat{f}_{BC} = f_{BC} - f_B - f_C + f_0, \quad (16)$$

$$\hat{f}_{ABC} = f_{ABC} - f_{AB} - f_{AC} - f_{BC} + f_A + f_B + f_C - f_0. \quad (17)$$

In the base case, $\hat{f}_0$, the contribution from factors is not considered; $\hat{f}_A$ is the pure contribution of factor A while $\hat{f}_B$ and $\hat{f}_C$ are the pure contributions of factors B and C , respectively. Term $\hat{f}_{AB}$ is the contribution from pure interaction of factors A and B , and similarly for the other interaction terms.

5. Design of experiments: Main and interaction effects

Computing the effects of various factors is a commonly used method from the design of experiments toolkit. The effects are typically calculated from the results of a full factorial experiment. The *main effect* of a factor is the change in the response induced by a change in the factor's level (Wu and Hamada 2009; Montgomery 2013). A main effect is the difference between the average of all observations at the high level of the factor and the average at the low level. Using the notation of Wu and Hamada (2009), the main effect of a factor A is

$$ME(A) = \bar{z}(A +) - \bar{z}(A -), \quad (18)$$

where $\bar{z}(A +)$ is the average of all runs when A is present (high) and $\bar{z}(A -)$ is the average of all runs when A is absent (low).

Using the three-factor experiment from Table 1, the main effect of factor A is found by

$$\begin{aligned} ME(A) &= \frac{1}{4}(f_A + f_{AB} + f_{AC} + f_{ABC}) - \frac{1}{4}(f_0 + f_B + f_C + f_{BC}) \\ &= \frac{1}{4}(f_A + f_{AB} + f_{AC} + f_{ABC} - f_0 - f_B - f_C - f_{BC}). \end{aligned} \quad (19)$$

The other main effects are similarly calculated.

When the response for a factor changes as the levels of or more other factors change, we say there is an *interaction effect*. Once again using Wu and Hamada's notation, the interaction effect of factors A and B is

$$INT(A, B) = \frac{1}{2}[ME(A|B+) - ME(A|B-)], \quad (20)$$

where $ME(A|B+)$ is the *conditional main effect* of A when B is present and likewise $ME(A|B-)$ when B is absent. These are calculated by

$$ME(A|B+) = \bar{z}(A + |B +) - \bar{z}(A - |B +), \quad (21)$$

$$ME(A|B-) = \bar{z}(A + |B -) - \bar{z}(A - |B -), \quad (22)$$

where $\bar{z}(A + |B +)$ is the average of the runs with both A and B present and similarly for the other terms. The interaction of AB compares the effect of A when B is on and when B is off. If there is a significant difference in how A affects the output depending on B , the interaction effect will be large.

Using our example from Table 1, the interaction effect of A and B is

$$\begin{aligned} INT(A, B) &= \frac{1}{2} \left\{ \left[\frac{1}{2}(f_{AB} + f_{ABC}) - \frac{1}{2}(f_B + f_{BC}) \right] \right. \\ &\quad \left. - \left[\frac{1}{2}(f_A + f_{AC}) - \frac{1}{2}(f_C + f_0) \right] \right\} \\ &= \frac{1}{4}(f_{AB} + f_{ABC} - f_B - f_{BC} - f_A - f_{AC} + f_C + f_0) \end{aligned} \quad (23)$$

and similarly for the interactions of A and C and of B and C .

The process is similar for interaction effects involving three or more factors.

6. Coding, simple effects, and factor separation

To understand the relationship between coding, effects, and factor separation, consider the *general linear model* used in multiple linear regression. The response y is a function of factors $x_1, x_2, \dots, x_m$ such that

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon, \quad (24)$$

where ε is random noise assumed to be normally distributed with mean 0 and variation σ^2 (Wu and Hamada 2009).

We can take the expected value of Eq. (24) to find that

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m. \quad (25)$$

We account for interaction among the x_i in Eq. (24) by using $\beta_{ij\dots}$ with $x_i x_j \dots$. For example, in the two-factor case:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon. \quad (26)$$

When the x_i correspond to categorical instead of numerical variables, x_i are typically coded. One option employs *effects coding* by using -1 to denote when a factor is absent and 1 for present while another employs *dummy coding* wherein 0 denotes a factor is absent and 1 its presence (Hardy 1993). We also note that the coefficients of dummy coded linear regression are known as *simple effects* as opposed to the main effects because they only measure the effect when all other factors are off due to the 0 coding (Collins et al. 2009, 2014).

Since regression coefficients in Eq. (24) measure a one-unit change and the main effects measure a two-unit change (-1 to 1) when using *effects coding*, the regression coefficients are one-half the main and interaction effects; thus, the coefficients quantify how the factor or combination of factors affects the average of all the runs, which is also known as the “grand mean” (Kugler et al. 2012; Wu and Hamada 2009; Montgomery 2013).

The interpretation of model in Eq. (24) is not as clear when using *dummy coding*, which we can see by constructing the equations for a factorial design. Let y_i be the run with just factor x_i present and y_{ijk} the run with factors x_i , x_j , and x_k present, and so on.

Then with dummy coding,

$$E(y_0) = \beta_0, \quad (27)$$

$$E(y_i) = \beta_0 + \beta_i, \quad (28)$$

$$E(y_{ij}) = \beta_0 + \beta_i + \beta_j + \beta_{ij}, \quad (29)$$

$$E(y_{123\dots n}) = \beta_0 + \sum_{i=1}^n \beta_i + \sum_{i,j=1,2}^{n-1,n} \beta_{ij} + \sum_{i,j,k=1,2,3}^{n-2,n-1,n} \beta_{ijk} + \dots + \beta_{123\dots n}. \quad (30)$$

A full factorial design for n factors with two levels and no replication would yield 2^n equations. The run $y_{ijk\dots}$ is used to estimate $E(y_{ijk\dots})$. With the 2^n runs from the full factorial, all 2^n β can be computed using

$$E(y_{123\dots n}) = f_{ijk\dots n}. \quad (31)$$

Notice this setup is identical to FS with different notation. Every $\hat{f}$ in Eq. (9) corresponds to a β in Eq. (30):

$$\hat{f}_{ij\dots} = \beta_{ij\dots}. \quad (32)$$

Consequently, FS effects are equivalent to multiple linear regression coefficients for a full factorial with one replication when using dummy coded variables. Dummy coding, however, can be used with any number of replications. For a full factorial design, the mean of the replications of each run is used to find the coefficients.

FS was only designed for a set of runs with a single replication, but simple effects can be calculated with or without replication. Several equivalent ways to express the simple effect of A are

$$\begin{aligned} \text{SE}(A) &= \text{ME}(A|\text{other factors} -) \\ &= \bar{z}(A|\text{other factors} -) - \bar{z}(A -|\text{other factors} -) \\ &= \bar{f}_A - \bar{f}_0, \end{aligned} \quad (33)$$

TABLE 2. The 2^2 factorial design with replication.

Run number	Factor A	Factor B	Simulation result average
1	—	—	$\bar{f}_0$
2	+	—	$\bar{f}_A$
3	—	+	$\bar{f}_B$
4	+	+	$\bar{f}_{AB}$

where $\bar{f}_A$ and $\bar{f}_0$ are the averages of the replications of f_A and f_0 , respectively, if the runs had been replicated. We see then that FS is the special case of calculating the simple effects when there is no replication.

7. Simple versus main effect models

DoE simple effects calculate the difference in the response between when A is present and when A is absent, all other factors absent, or the interaction effect of A and another factor, all the other factors absent (Collins et al. 2009). On the other hand, DoE main effects calculate the difference between when A is present and when A is absent, averaging over all levels of the other factors. The main effect answers questions about the average effect of A or the average interaction effect of A and another factor combined (Collins et al. 2009).

As demonstrated in the previous section, simple effects are identical to the coefficients computed from coding multiple linear regression with 0 and 1 dummy coding (Kugler et al. 2012). Main and interaction effects are twice the coefficients computed when coding multiple linear regression using -1 and 1 effects coding (Kugler et al. 2012). Even though the coefficients from these two coding methods are different, the predictions are identical because the coding simply scales the variables (Hardy 1993).

To demonstrate this, let us use the effects from the two methods to create models to predict the output for different configurations. We can do this by creating the linear regression equations for each coding scheme. For simplification, we will show this for the two factor situation. We will assume replication is present because without replication both models simply reproduce the data. The FS would be performed on the mean of the replications of each run. Table 2 gives the averaged runs for the experiment, and Table 3 shows the computations for the effect estimates.

The simple effects model is created from the factor separation effects:

$$\text{FS}(x_1, x_2) = \hat{f}_0 + \hat{f}_1 x_1 + \hat{f}_2 x_2 + \hat{f}_{12} x_1 x_2, \quad 0 \leq x_1, x_2 \leq 1. \quad (34)$$

We construct the main effects model using w_i instead of x_i because of the difference in coding. $\bar{z}$ is the average

TABLE 3. The 2² factorial effects.

Effect	Symbol	Result
Base case	$\hat{f}_0$	$\bar{f}_0$
Contribution of A	$\hat{f}_A$	$\bar{f}_A - \bar{f}_0$
Contribution of B	$\hat{f}_B$	$\bar{f}_B - \bar{f}_0$
Contribution of interaction of A and B	$\hat{f}_{AB}$	$\bar{f}_{AB} - \bar{f}_A - \bar{f}_B + \bar{f}_0$
Grand mean	$\bar{z}$	$(1/4)(\bar{f}_0 + \bar{f}_A + \bar{f}_B + \bar{f}_{AB})$
Main effect of A	ME(A)	$(1/2)(\bar{f}_A + \bar{f}_{AB} - \bar{f}_B - \bar{f}_0)$
Main effect of B	ME(B)	$(1/2)(\bar{f}_B + \bar{f}_{AB} - \bar{f}_A - \bar{f}_0)$
Interaction effect of A and B	INT(AB)	$(1/2)(\bar{f}_{AB} - \bar{f}_A - \bar{f}_B + \bar{f}_0)$

of all the runs, known as the “grand mean.” The main effects model for two variables is

$$\text{DOE}(w_1, w_2) = \bar{z} + \frac{\text{ME}(A)}{2}w_1 + \frac{\text{ME}(B)}{2}w_2 + \frac{\text{INT}(A, B)}{2}w_1w_2, \quad -1 \leq w_1, w_2 \leq 1. \quad (35)$$

Notice that we can convert between x_i and w_i :

$$w_1 = 2x_1 - 1, \quad (36)$$

$$w_2 = 2x_2 - 1. \quad (37)$$

By substituting w_1 and w_2 into the DoE model we find

$$\text{DOE}(x_1, x_2) = \bar{z} + \frac{\text{ME}(A)}{2}(2x_1 - 1) + \frac{\text{ME}(B)}{2}(2x_2 - 1) + \frac{\text{INT}(A, B)}{2}(2x_1 - 1)(2x_2 - 1). \quad (38)$$

Substituting the equations for ME(A), ME(B), INT(A, B), and $\bar{z}$ into Eq. (38) and simplifying yields

$$\text{DOE}(x_1, x_2) = \bar{f}_0 - \bar{f}_0x_1 - \bar{f}_0x_2 + \bar{f}_Ax_1 + \bar{f}_Bx_2 + \bar{f}_0x_1x_2 - \bar{f}_Ax_1x_2 - \bar{f}_Bx_1x_2 + \bar{f}_{AB}x_1x_2. \quad (39)$$

Similarly, substituting the equations for $\hat{f}_0, \hat{f}_1, \hat{f}_2$, and $\hat{f}_{12}$ into the factor separation model in Eq. (34) gives us

$$\text{FS}(x_1, x_2) = \bar{f}_0 - \bar{f}_0x_1 - \bar{f}_0x_2 + \bar{f}_Ax_1 + \bar{f}_Bx_2 + \bar{f}_0x_1x_2 - \bar{f}_Ax_1x_2 - \bar{f}_Bx_1x_2 + \bar{f}_{AB}x_1x_2. \quad (40)$$

Notice that

$$\text{DOE}(x_1, x_2) = \text{FS}(x_1, x_2). \quad (41)$$

Thus, both models will predict the same output for any given input. While this proof is only for two variables, it is known that effects coding and dummy coding always create identical predictions for any number of variables.

The only difference is in interpreting the coefficients (Kugler et al. 2012). In fact Hardy (1993) puts it well: “The different coding scheme affects the way the information is captured—the manner in which group differences are arrayed—but it does not affect the overall picture because the underlying structure remains unchanged from earlier estimations; we simply view it from a different angle.”

8. Shallow-water equations example

We demonstrate FS and DoE main effects using a shallow-water equations model coded in Python by Paul Connolly that we modified for our use. The equations were solved using the Lax–Wendroff method (Connolly 2018). This shallow-water model allows for varying different parameters. We investigate how the different parameters affect the simulation output. In particular, for this experiment our goal was to understand how the simulated surface height changed with varying wind speed and latitude.

The height field was initialized so the initial wind was uniform westerly in the northern half of the domain and uniform easterly in the southern half of the domain, creating a sharp shear. Since in Stein and Alpert’s (1993) example they varied two parameters, surface fluxes and terrain, we likewise varied two parameters: latitude (factor A) and wind speed (factor B). We placed the domain at a latitude of 20° as the low setting and moved it to a latitude of 45° as the high. We used 20 m s^{−1} for the low wind speed and 50 m s^{−1} for the high wind speed. The surface height at 30 h into the simulation, before the

TABLE 4. The four shallow-water model simulations.

Run number	Latitude (°N) (factor A)	Wind speed (m s ^{−1}) (factor B)	Surface height (m) (response)
1	20	20	f_0
2	45	20	f_A
3	20	50	f_B
4	45	50	f_{AB}

FIG. 1. SWE simulation result run 1: All factors low.

FIG. 2. SWE simulation result run 2: Wind speed low, latitude high.

point when the sharp shear starts to become unstable, was used as the response variable.

Following a full factorial design, the four simulation runs are found in Table 4. The results are in Figs. 1–4.

We then calculated the DoE main effects (Figs. 5–8) and FS effects (Figs. 9–12) at every point in the domain and compared the results. *Note that the scales differ between*

the grand mean or base effect figures and the remaining effect figures.

Comparing DoE main effects and FS effects, it is clear that the results are similar but not the same. Figures 5 and 9 are the constants for the regression equations. DoE main effects use the grand mean in Fig. 5 as the constant and compare changes in factor levels to this

FIG. 3. SWE simulation result run 3: Wind speed high, latitude low.

FIG. 4. SWE simulation result run 4: Wind speed high, latitude high.

FIG. 5. Design of experiments: Grand mean of all SWE runs.

FIG. 6. Design of experiments: Main effect of higher latitude.

mean while FS uses the base case, no factor run, in Fig. 9 as the constant and compares changes to this base.

Figures 6 and 10 both show the effect of changing the latitude to 45° . While they have the same general shape, the main effect of latitude in Fig. 6 has a steeper slope than the pure contribution of latitude in Fig. 10. This is because the main effect calculates the effect of latitude at both low and high wind speeds while FS

calculates the effect of latitude just at low wind speeds. High wind speeds at high latitudes increase the slope, which is shown in the main effect of latitude. Notice the FS interaction in Fig. 12 has a steeper slope than the DoE main effects interaction in Fig. 8 so the information that higher latitude and higher wind speeds increase the slope is contained in the FS interaction in Fig. 12.

FIG. 7. Design of experiments: Main effect of higher wind speed.

FIG. 8. Design of experiments: Interaction effect of higher wind speed and latitude.

FIG. 9. Factor separation: Zero-state low latitude, low wind speed.

FIG. 10. Factor separation: Contribution (simple effect) of higher latitude.

Thus, both methods tell us that in the SWE simulations the factors each individually increase the slope as well as interact together to increase the slope as the factors increase.

9. Advantages of DoE for factor separation

Alpert and Sholokhman (2011b) compared FS effects to main effects, which they called “factorial modeling,”

and concluded FS works better for atmospheric simulations. They did identify a few areas of difficulty with FS, especially that of the exponentially increasing computational cost. Here we address their assertions of the benefits of FS over main effects and demonstrate how DoE uses fractional factorial designs with smaller numbers of runs to decrease the computational cost.

FIG. 11. Factor separation: Contribution (simple effect) of higher wind speed.

FIG. 12. Factor separation: Synergy of higher wind speed and latitude.

We briefly discuss the benefits of DoE for factors with more than two levels.

Alpert and Sholokhman (2011b) presented two main advantages of FS effects over DoE effects.

- 1) The sum of all FS effects equals the full run results.
- 2) The zero state, also called basic state, is one of the FS effects.

Stein and Alpert pointed out in their original paper, and then reiterated it in the book that followed, that all 2^n FS pure contributions and synergies add up to the “full run” (all factors on), which they also called the *control run* (Stein and Alpert 1993; Alpert and Sholokhman 2011a). They assert this allows one to view the effects as a percentage of the full run. For example, using simple numbers in a two-factor experiment, if the pure contribution of a factor A is $\hat{f}_A = 2$ and the full run result is $f_{AB} = 8$, then the pure contribution of A is 25% of the full run. This can be helpful, but it only applies to the control run. It would be a mistake to then say A contributes 25% of the result of any run including A . Thus, the usefulness of the percentage-wise analysis is limited.

They also discussed the importance of the zero-state run with all factors off and pointed out, “It is shown here and in many earlier FS studies that in the atmosphere the zero-state contribution can be significant and should be calculated” (Alpert and Sholokhman 2011b). Because the main and interaction effects do not calculate this zero state, Alpert and Sholokhman assert FS is a better method. It is true the zero state is not calculated as a main effect, but it is always one of the runs in a full factorial model so that information still exists when using DoE.

The biggest difference between FS and DoE is FS effects are compared to the zero state while main effects are compared to the grand mean. It is easy to see that, depending on the experiment, sometimes it is more helpful to compare the effects to the grand mean and other times to compare them to the zero state. It depends on the purpose of the particular experiment.

When a full factorial design can be used, both types of effects are easily computed. In fact, using both methods allows an even better glimpse into how the factors contribute to the results. As Alpert and Sholokhman (2011b) mention, however, a full factorial design can have a high computational cost as the number of factors grows while leaving a factor out of the analysis can have unintended consequences since synergistic contributions of the factors will then be wrapped into the other contributions (Alpert and Sholokhman 2011b). One solution is to not use a full factorial design. Because oftentimes the higher-order synergies become small, they may be neglected. A full factorial is not necessary in order

to calculate the first- and second-order contributions alone (Alpert and Sholokhman 2011b).

This concept that the higher-order synergies may not be important is known in the design of experiments as the effect hierarchy principle. This principle is stated in Wu and Hamada (2009) as follows:

- “Lower order effects are more likely to be important than higher order effects.”
- “Effects of the same order are equally likely to be important.”

This principle does not always hold, but it is useful for experiments with few runs but a large number of factors (Wu and Hamada 2009).

The design of experiments deals extensively with the problem of not being able to run a full factorial experiment. This is addressed in the next section.

a. Fractional factorial designs and aliasing

Design of experiments offers *fractional factorial designs*, also known as 2^{k-p} designs, that use 2^{k-p} runs instead of 2^k runs for k factors (Wu and Hamada 2009). These can significantly decrease the computational cost of the design, but the consequence of this is a concept known as *aliasing*. Aliasing of effects occurs when the data cannot distinguish the estimate of one effect from the estimate of another effect (Wu and Hamada 2009). An example of a fractional factorial design and the resulting aliased effects can be found in the appendix.

Note fractional factorial designs should not be confused with the concept of *fractional treatment* as proposed by Krichak and Alpert (2002). Fractional treatments deal with varying the intensities of the factors, which is not what is being discussed here.

The trade-off for fewer runs is aliasing, but there are ways to still gather information from aliased effects. Because of the effect hierarchy principle, the lower-order effects are more likely to be significant than the higher-order ones, so the higher-order interaction of an alias in many cases may be ignored (Wu and Hamada 2009). Previous knowledge of effect importance can also help in sorting through the aliased effects.

“Dealiasing” effects is a current area of study in the design of experiments where inferences about the effects can still be made from aliased effects. While Wu (2018) admits, “When the experiment does not contain enough information, one cannot squeeze too much out of it,” strides are being made in the ability to dealias effects. It is out of the scope of this paper to go into detail of the current techniques for dealiasing, but one relevant emerging technique is that of using the conditional main effect (CME) when the main effect is aliased (Su and Wu 2017; Mak and Wu 2019; Peng 2018). CMEs are actually similar, and in

TABLE A1. Resolution IV 2^{4-1} fractional design.

Run	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ACD	ABD	BCD	ABCD
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	1	-1	1	1
3	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	-1	1	1	1
4	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1
5	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1
6	-1	1	-1	1	-1	1	-1	-1	1	-1	1	1	-1	-1	1
7	-1	-1	1	1	1	-1	-1	-1	-1	1	1	-1	1	-1	1
8	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1

some cases identical, to FS effects. They are identical when the CME is conditioned on all other factors being zero. The CME method trades out aliased effects with CMEs. This combination of using both traditional effects and conditional main effects seems promising (e.g., Yoshida 2018).

b. Simple effects of fractional designs

It was considered whether or not FS effects, or simple effects, can be computed from the results of a fractional experiment. While we do not have a complete answer at this time, we know it is not always possible to perform FS as described by Stein and Alpert if there are missing runs. An advantage of using main effects is it is always possible to calculate them. For example, if the run with just factor *A* is not completed, then FS cannot be performed to find the pure contribution of *A*. The main effect of *A*, however, can still be computed using the other runs, though it may be aliased with another effect.

c. Factors with more than two levels

One last advantage of using DoE to approach such experiments is there are methods for handling factors with more than two levels (Wu and Hamada 2009). Krichak and Alpert (2002) only briefly address the idea of exploring values in between the high and low levels. Since few factors are as simple as on and off, the ability to incorporate more levels is helpful.

10. Conclusions

The ability to tease out factor effects and interactions from a simulation is invaluable for many disciplines, and the benefits of factor separation for the atmospheric sciences are numerous. We showed here that FS effects are the same as simple effects, a tool used in the design of experiments. FS and DoE main effects both correspond to multiple linear regression coefficients with different coding schemes, and when used to create models, the two methods produce identical output (Kugler et al. 2012). The interpretation of the simple effects, or FS effects, is different from that of the main effects, but in general, as demonstrated with the shallow-water equations,

we can come to the same conclusions using either method (Hardy 1993).

In the case of a small number of factors with two levels, where a full factorial design is possible, neither method is inherently better, but one way may be better suited to the specific situation and questions being asked. It is critical, however, that whichever method is used is made clear since the interpretations are different.

We recommend continued investigation into how the design of experiments can benefit the atmospheric sciences beyond the FS method since the DoE approach includes fractional factorial designs as well as the ability to account for more than two levels, allowing for greater applicability to atmospheric simulations.

Acknowledgments. The authors wish to acknowledge Dr. Ligia Bernardet of NOAA/GSD, the lead for DTC's model test bed, who asked one of us (Smith) how DoE related to factor separation at the 2018 Annual AMS meeting in response to our paper presented therein. Her question sparked this research. Research was sponsored by the U.S. CCDC Army Research Laboratory and was accomplished under Cooperative Agreement W911NF-18-2-0252. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. CCDC Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

APPENDIX

A Fractional Factorial Design Example

As an example of a fractional design, Table A1 is a one-half fraction design for four factors, *A*, *B*, *C*, and *D*. A full factorial for four factors would require $2^4 = 16$ runs, but this design has $2^{4-1} = 8$ runs.

In the design matrix, there are four columns for the main effects, six columns for the two-factor interactions,

four columns for the three-factor interactions, and one column for the four-factor interaction. These are generated by multiplying the corresponding A , B , C , and D columns. For example, column AB is generated by multiplying pairwise the entries for A and B in each row. This is a simple way to see which effects are aliased. If the columns for two effects are identical, the effects are aliased. We denote aliased effects by $A = BCD$ to show column A and column BCD are identical and thus $ME(A)$ and $INT(BCD)$ are aliased. The aliased effects for the design in Table A1 are as follows:

$$A = BCD, \quad (A1)$$

$$B = ACD, \quad (A2)$$

$$C = ABD, \quad (A3)$$

$$D = ABC, \quad (A4)$$

$$AB = CD, \quad (A5)$$

$$AC = BD, \quad (A6)$$

$$AD = BC, \quad (A7)$$

$$I = ABCD, \quad (A8)$$

where $I = ABCD$, the column with all ones, is known as the “defining relation.” This type of design is a Resolution IV design because of the four letters in the defining relation (Wu and Hamada 2009).

The main effects are each aliased with a three-way interaction effect. Using the effect hierarchy principle, if we assume the three-way interactions are zero, we can estimate the main effects. The two-way effects, however, cannot be estimated because they are all aliased with another two-way interaction.

REFERENCES

- Alpert, P., 2011: Meso-meteorology: Factor separation examples in atmospheric meso-scale motions. *Factor Separation in the Atmosphere: Applications and Future Prospects*, P. Alpert and T. Sholokhman, Eds., Cambridge University Press, 53–66.
- , and T. Sholokhman, Eds., 2011a: *Factor Separation in the Atmosphere: Applications and Future Prospects*. Cambridge University Press, 274 pp.
- , and —, 2011b: Some difficulties and prospects. *Factor Separation in the Atmosphere: Applications and Future Prospects*, P. Alpert and T. Sholokhman, Eds., Cambridge University Press, 237–244.
- Berger, A., M. Claussen, and Q. Yin, 2011: Factor separation methodology and paleoclimates. *Factor Separation in the Atmosphere: Applications and Future Prospects*, P. Alpert and T. Sholokhman, Eds., Cambridge University Press, 28–52.
- Collins, L. M., J. J. Dziak, and R. Z. Li, 2009: Design of experiments with multiple independent variables: A resource management perspective on complete and reduced factorial designs. *Psychol. Methods*, **14**, 202–224, <https://doi.org/10.1037/a0015826>.
- , —, K. C. Kugler, and J. B. Trail, 2014: Factorial experiments: Efficient tools for evaluation of intervention components. *Amer. J. Prev. Med.*, **47**, 498–504, <https://doi.org/10.1016/j.amepre.2014.06.021>.
- Connolly, P. J., 2018: Shallow water practice model, version 1.0.0. Zenodo, <https://doi.org/10.5281/zenodo.1478060>.
- Fisher, R. A., 1971: *The Design of Experiments*. 8th ed. Hafner Publishing Company, 248 pp.
- Hardy, M. A., 1993: Regression with dummy variables. *Quantitative Applications in the Social Sciences*, SAGE Publications, 90 pp.
- Krichak, S. O., and P. Alpert, 2002: A fractional approach to the factor separation method. *J. Atmos. Sci.*, **59**, 2243–2252, [https://doi.org/10.1175/1520-0469\(2002\)059<2243:AFATTF>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<2243:AFATTF>2.0.CO;2).
- Kugler, K. C., J. B. Trail, J. J. Dziak, and L. M. Collins, 2012: Effect coding versus dummy coding in analysis of data from factorial experiments. The Pennsylvania State University Tech. Rep., <http://methodology.psu.edu/media/techreports/12-120.pdf>.
- Mak, S., and C. F. J. Wu, 2019: cmenet: A new method for bi-level variable selection of conditional main effects. *J. Amer. Stat. Assoc.*, **114**, 844–856, <https://doi.org/10.1080/01621459.2018.1448828>.
- Montgomery, D. C., 2013: *Design and Analysis of Experiments*. John Wiley and Sons, 730 pp.
- National Research Council, 1995: *Statistical Methods for Testing and Evaluating Defense Systems: Interim Report*. National Academies Press, 84 pp., <https://doi.org/10.17226/9074>.
- Peng, C.-Y., 2018: Discussion. *Ann. Inst. Stat. Math.*, **70**, 269–274, <https://doi.org/10.1007/s10463-017-0640-y>.
- Reuter, G. W., 2011: Application of the factor separation methodology to quantify the effect of waste heat, vapor and pollution on cumulus convection. *Factor Separation in the Atmosphere: Applications and Future Prospects*, P. Alpert and T. Sholokhman, Eds., Cambridge University Press, 163–170.
- Smith, J. A., and R. S. Penc, 2016: A design of experiments approach to evaluating parameterization schemes for numerical weather prediction: Problem definition and proposed solution approach. *Conf. on Applied Statistics in Defense*, Fairfax, VA, Interface Foundation of North America and George Mason University College of Science, 4183–4192.
- , —, and J. W. Raby, 2018: Statistical design of experiments in numerical weather prediction: Emerging results. *25th Conf. on Probability and Statistics*, Austin, TX, Amer. Meteor. Soc., 6.1, <https://ams.confex.com/ams/98Annual/meetingapp.cgi/Paper/326537>.
- , —, —, and J. L. Cleveland, 2019: Some conclusions on applying statistical design of experiments to numerical weather prediction. *18th Conf. on Artificial and Computational Intelligence and its Applications to the Environmental Sciences*, Phoenix, AZ, Amer. Meteor. Soc., TJ17.4, <https://ams.confex.com/ams/2019Annual/meetingapp.cgi/Paper/352596>.
- Stein, U., and P. Alpert, 1993: Factor separation in numerical simulations. *J. Atmos. Sci.*, **50**, 2107–2115, [https://doi.org/10.1175/1520-0469\(1993\)050<2107:FSINS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<2107:FSINS>2.0.CO;2).
- Su, H., and C. F. J. Wu, 2017: CME analysis: A new method for unraveling aliased effects in two-level fractional factorial experiments. *J. Qual. Technol.*, **49**, 1–10, <https://doi.org/10.1080/00224065.2017.11918181>.

- Thunis, P., and Coauthors, 2019: Source apportionment to support air quality planning: Strengths and weaknesses of existing approaches. *Environ. Int.*, **130**, 104825, <https://doi.org/10.1016/j.envint.2019.05.019>.
- Waugh, D. W., A. M. Hogg, P. Spence, M. H. England, and T. W. N. Haine, 2019: Response of Southern Ocean ventilation to changes in midlatitude westerly winds. *J. Climate*, **32**, 5345–5361, <https://doi.org/10.1175/JCLI-D-19-0039.1>.
- Wu, C. F. J., 2018: Rejoinder. *Ann. Inst. Stat. Math.*, **70**, 279–281, <https://doi.org/10.1007/s10463-017-0639-4>.
- , and M. S. Hamada, 2009: *Experiments Planning, Analysis, and Optimization*. 2nd ed. Wiley Series in Probability and Statistics, John Wiley and Sons, 716 pp.
- Yang, L., J. Smith, and D. Niyogi, 2019: Urban impacts on extreme monsoon rainfall and flooding in complex terrain. *Geophys. Res. Lett.*, **46**, 5918–5927, <https://doi.org/10.1029/2019GL083363>.
- Yoshida, R., 2018: Discussion on the paper by Professor Wu. *Ann. Inst. Stat. Math.*, **70**, 275–278, <https://doi.org/10.1007/s10463-017-0641-x>.