A Design of Experiments Approach to Evaluating Parameterization Schemes for Numerical Weather Prediction: Problem Definition and Proposed Solution Approach

Jeffrey A. Smith[1], Richard S. Penc[1]
[1]U.S. Army Research Laboratory, RDRL-CIE-M, WSMR, NM 88002-5501

Abstract
Numerical Weather Prediction (NWP) is the science of forecasting weather or climatic conditions based on past and present observations using computational methods applied to mathematical representations of the atmosphere. Temporally, weather forecasts range from a few hours to a several days in the future, while climate forecasts range from several months to years (or decades) into the future. Spatially, forecasts can cover small scale, highly resolved "local" weather conditions to large scale global weather features and climates.

The foundation of NWP is the conservation of mass, heat, momentum, and water vapor, along with other gaseous and aerosol materials over a region of interest called the domain (Pielke 2002; Warner 2011). The conservation equations are nonlinear, partial-differential equations that are nearly impossible to solve analytically except in a few ideal cases. Practical solution approaches for these equations employ numerical methods to obtain approximate forecasts for a domain represented by a finite and generally regular set of discrete "grid points". Discretizing the domain means that atmospheric processes occuring at sub-grid scales cannot be resolved by the modeled physics; however, these unresolved effects must be accounted for to maintain conservation. Such accounting is done via parameterizations that address physical effects (terrain, land use, turbulence, moisture, etc.) which occur at sub-grid scales. Depending on which parts of the atmosphere researchers consider, there are a number of parametric approaches to model these physical effects. It is difficult to efficiently explore how these parameterizations interact over a domain to produce a forecast; however, we require this knowledge to conduct trade-off studies and inform the selection of parameterization schemes to make the NWP robust for a variety of weather conditions.

Statistical design of experiments, a technique applied successfully in other areas to large scale simulation models, shows promise in assisting in a structured exploration of these parameterized processes in NWP codes. In this article, we develop an extended problem definition; we present a method for developing a design matrix suitable for that problem; and, we illustrate how to apply that design to study the role parameterizations play in a relevant forecast metric of interest.

Key Words: numerical weather prediction, forecast, weather, climate, model, design of experiments, DoE, experiment design, Latin Hypercube, design matrix, orthogonal arrays, mixed designs, numerical weather prediction, NWP, weather research and forecasting, WRF, WRF-ARW

1. Introduction

From the early work of Bjerknes and Richardson (Lynch 2006), Charney's (1948, 1949) work on the numerical prediction in the atmosphere, and Smagorinsky's (1963) landmark paper on the development of a global circulation model, our ability to forecast weather and climatic conditions has grown ever more more capable. Yet at the same time, these models, with all their attendant assumptions and approximations together with the computational challenges arising from ever finer modeling resolutions, can interact in a number of ways that introduce even more uncertainty into the forecast (Stevens and Bony 2013). This complexity challenges those who rely on these models to support decision making (Vecchi and Villarini 2014) and to assess regional impacts (Hall 2014; Schindler and Hilborn 2015) despite ever increasing amounts of data (Overpeck et al. 2011). Some of this inherent uncertainty can be attributed to how these models introduce parametric implementations of sub-grid physics effects called "parameterizations" (Stensrud 2007). Yet, many of the parameterizations, while studied in varying levels of detail by their developers, have not been studied in detail for their role in producing forecasts. It is in this later aspect that we see a role for experimental design.

Experimental design reaches back to the work of Fisher (1935) who developed the basic techniques for agricultural science. Other researchers (e.g., Box et al. 1978; Montgomery 1997; and Deming 2000) extended Fisher's work into areas such as industrial process understanding and control. Recently, with advances in computational capability, researchers, e.g., Kleijnen et al. (2005), and Kleijnen (2008), have applied experimental design to the study of complex simulation codes. In approximently 80 years of research, experimental design has evolved into a robust and comprehensive collection of methodologies that allow rigorous experimentation in many complex systems far removed from Fisher's initial application. Yet despite the evolution of these techniques there is little direct evidence to suggest that researchers have emplied these techniques to study numerical weather prediction (NWP). Absent any direct evidence we note that some (Berci et al. 2014; Rahimi et al. 2014; or Zhu et al. 2015) have applied experimental design methods to computational fluid dynamics codes as part of an engineering development process. We also recognize that computational fluid dynamics codes share many of the same complexities exhibited by NWP, a fact that suggests that experimental design may prove useful in weather forecasting and the analysis of the attendant models.

This article serves as an extended, and more complete definition, of a problem presented at a clinical session held at the recent Conference on Applied Statistics in Defense (Smith and Penc 2015a). In this session, we expressed our initial thoughts on the exploration of an existing NWP code via modern experimental design techniques. In Section 2, we present a process model for a generic NWP code along with an interaction diagram of parameterizations for a specific NWP model. In Section 3, we present methods that allow sampling of that process model. In Section 4, we use an approach developed at the Naval Postgraduate School, Monterey, CA to create an initial design matrix for our problem. In Section 4, we also complete the problem statement by proposing how we can apply design of experiments to assessing how these various parameterizations interact to produce a "skillful" forecast. In Section 5, we summarize our efforts to explore NWP with design of experiments.

2. Problem

Since the atmosphere is a fluid, each approach to NWP uses some form of the Navier-Stokes equations of fluid and thermodynamics together with samples of the current atmospheric state taken up to a given time to estimate its future state. The Navier-Stokes equations are nonlinear partial-differential equations that are nearly impossible to solve analytically except in a few ideal cases. Therefore, practical implementations of NWP employ numerical methods to obtain approximate solutions at a set of finite, and generally regular, collection of of discrete “grid points” called a domain (Pielke 2002; Warner 2011).

2.1 A Process Model for NWP

Observing that a forecast takes input conditions up to some current time and maps them to a set of conditions at some future time via a suitably chosen mechanism, we model a forecast mathematically as:

$$M: I_\tau \to O_t \tag{1}$$

where I_τ denotes the set of inputs up to time τ with zero being the arbitrary start time; O_t denotes the forecast at some time t assumed to be greater than τ; M the mechanism that produces the forecast; and the right arrow signifies ‘maps’. For this article, we choose the term mechanism to be synonymous with the equations of fluid and thermodynamics implemented in a form that can be solved numerically on a digital computer — the NWP code. Though frequently used interchangeably, we forego the use of the terms “model” and “simulation” in favor of the term mechanism to avoid the often overloaded meanings ascribed to “model” and “simulation” by various communities.

Regardless of the NWP application, whether it be for a weather forecast or a climate prediction, the mechanism must be initialized. The process of entering observed data into the mechanism as an estimate of the current atmospheric state, and establishing the lateral boundary conditions if the forecast is for a region, is called initialization. The set of inputs, I_τ, is taken to mean the meteorological conditions, e.g., temperature, humidity, wind speed and direction, etc., that initialize the mechanism M to produce the forecast. We draw this distinction to make clear the point that parameterizations are conditions placed upon the mechanism producing the forecast rather than inputs in the sense of observations of the atmospheric state; although, in the simulation sense of inputs parameterizations are configurable inputs. We denote the finite set of parameterizations classes as θ_k. Here each class represents specific implementations associated with a single type of sub-grid process.

Because M is a finite approximation (both temporally and spatially) of a continuous physical process, it introduces numerical errors that grow with time (Lorenz 1963). Additionally, I_τ represents measurements of the atmospheric state that also come with associated, and sometimes unknown, measurement errors. Consequently, we conclude that the forecast O_t is a random variable. For purposes of this article, we define a forecast, under a specific $\varphi \in \theta_k$, as:

$$M(\varphi): \widetilde{I_\tau} \to \widetilde{O_t} \tag{2}$$

where the various terms have just been defined and the ‘~’ denotes that these are random variables. While it is tempting to assert that M is a random function owing to the presence of "bugs" in the software, we expect that M is deterministic (hence bug free) because of a continuous development, testing, and maintenance process.

2.2 Parameterizations in a Particular NWP

The discussion thus far has been about generic NWP; however, our interest is in a community developed NWP modeling system known as the Weather Research and Forecasting (WRF) model, and specifically the Advanced Research WRF (ARW) version (Skamarock et al. 2008), hereinafter WRF-ARW. For our purposes, WRF-ARW is employed to downscale global forecasts to a more highly resolved forecast for a limited domain or region in space. For this article, the specific location and domain are unimportant. The WRF-ARW simulation core itself, absent the atmospheric state initialization data, serves as our mechanism of study.

There are 7-broad classes of physics parameterizations from which a user selects using a "namelist" to configure WRF-ARW. Within each parameterization class there are a number of options to choose from to model a particular physical process. To produce a forecast, a user selects one option per parameterization class at start time via the namelist which configures WRF-ARW. These configuration options are given in Table 1. We use $\varphi \in \theta_k$ to denote a $1\times k$ vector of selections chosen as namelist inputs (one choice per each of k classes) from Table 1.

Table 1: Cross reference of physics parameterizations to namelist entry along with physical process represented and number of available options.

Parameterization	Namelist Entry	Physical Process	Options
Microphysics	mp_physics	Moisture Transport	17
Long wave radiation	ra_lw_physics	Long wave solar radiation	6
Short wave radiation	ra_sw_physics	Shortwave solar radiation	6
Surface layer physics	sf_sfclay_physics	Near earth effects	9
Surface physics	sf_surface_physics	Land/atmosphere interface	8
Planetary boundary layer (PBL)	bl_pbl_physics	Turbulent atmospheric layer near earth	12
Cumulus physics	cu_physics	Clouds	6

As the mechanism $M(\varphi)$ executes (in this case WRF-ARW configured with a specific parameterizations) the schemes for each parameterization class interact at predescribed intervals (also a namelist input) such that, when coupled with the integration of the fluid and thermodynamic equations, we obtain a forecast from a given set of input conditions. These interactions are depicted in Figure 1, along with some of the physical processes produced such as rain. Also shown are some of the model variables that correspond to the specific form of the Navier-Stokes equations implemented in the code. Note that in Figure 1, the box labeled "Radiation" handles both the "Long wave" and "Short wave" radiation schemes identified in Table 1. Likewise, the box labeled "Surface" handles both the "Surface layer" and "Surface physics" entries. How parameterizations interact is our research interest in applying statistical design of experiments to the mechanism M of equation 2.

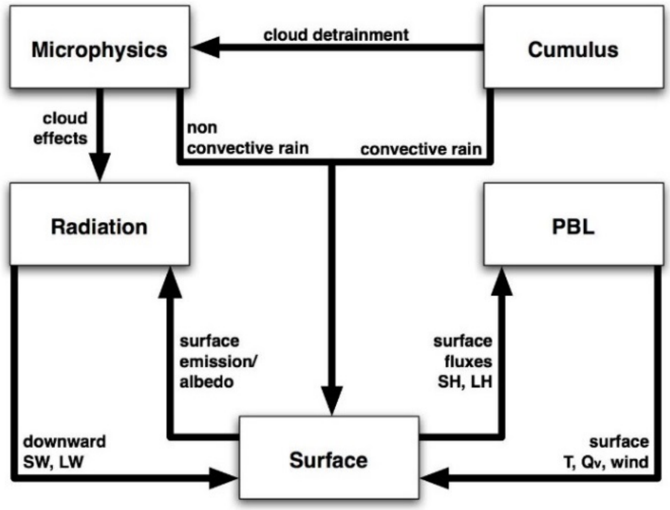

Figure 1: Direct interactions of parameterizations in WRF-ARW (Dudhia 2015).

3. Experimental Design

Montgomery (1997, p. 1) defines an experiment as "test or series of tests in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes in the output response." This definition holds regardless of whether one studies a physical system such as crops in the field or in a simulation code. Our objective is to design an experiment that enables us to extract as much information as we can, as efficiently as possible, from sampling the output of the mechanism at suitably chosen settings drawn from θ_k. We are mindful of the typical role experimental design plays in simulation analysis which is forming a meta-model of the simulation response. However, it is worth noting a principle that Cioppa and Lucas (2007) ascribe to Santner et al. (2003) for selecting designs that "allow one to fit a variety of models and provide information about all portions of the experimental region." By remembering this principle and reflecting it in our designs, we can add to our data set as our exploration progresses, and make the appropriate choice of meta-model as the data reveals it without relying upon meta-model specific assumptions *a priori* to create our design.

The barrier to applying design of experiments (DoE) to simulations rests largely on the fact that much of the theory behind DoE was developed to study real world applications and then adapted to the simulation world (Kleijnen et al. 2005). However, recent years have seen significant advances by a number of researchers (among them McKay et al. 1979; Sacks et al. 1989a; Sacks et al. 1989b; Law and Kelton 2000; Kleijnen et al. 2005; and Kleijnen 2008) in the development and cataloging of approaches and designs for simulation experiments, for example the SEED Center for Data Farming at the Naval Postgraduate School (2016) in Monterey, CA. Section 3 draws on the work done at the SEED Center and applies these approaches to numerical weather prediction codes.

3.1 Initial Attempts at an Experimental Design

A design matrix is a $n \times k$ matrix of n design points (vice runs) taken at various settings called levels that we denote as l for each of k factors. Taking WRF-ARW as the mechanism M (eq. 2) we wish to study, and inspecting both the documentation (Skamarock et al. 2008) and Table 1 reveal that each of the k parameterization classes (θ_k) to be categorical factors, and each scheme within a class to be a categorical variable. In addition to the parameterizations, we note from the documentation that there are a number of other factors, largely quantitative, that are of potential interest; however, for this particular effort, we restrict ourselves to the purely categorical parameterizations identified for the k parameterized physical processes shown in Table 1.

The categorical nature of Table 1 suggests a design based on Latin Squares or Taguchi methods (Montgomery 1997). Treating each of the k parameterization classes as a factor would allow a Latin Square design to identify factor effects but not factor interactions; while a Taguchi approach would allow us to identify both factor and two-factor interactions, but no higher order interactions. It is clear from both Figure 1 and our knowledge of atmospheric physics that interactions between classes of parameterizations are likely; however, it is also reasonable to suppose that higher order, those beyond two factor, interactions exist so it appears neither design method is suitable for our purposes.

One approach that will allow us to explore a high dimensional, categorical design space is some form of a l^k factorial design where there are l levels per each of k factors, the simplest case being $l = 2$ (Montgomery 1997). For the $l = 2{,}3$ and 4 cases there are cataloged designs available; however, the downside to the factorial approach is that for many of the factors in Table 1, l is quite large, so to find an appropriate design we would need to generate designs and search through those designs using some heuristic to find one suitable for our needs. In our case, there are on the order of three million different factor-level combinations for the classes identified in Table 1 so a heuristic search through this space is uncertain to produce a suitable design. We need an approach that allows us to explore this huge space of combinations efficiently, yet at the same time recognizes the computational expense of creating an extensive run set. In other words, a method that limits the number of design points (n) that we need to draw a statistically valid conclusion about the impact of parameterizations on our forecast producing mechanism.

3.2 Would a Latin Hypercube Approach Work?

McKay et al. (1979) first applied Latin Hypercube sampling to a computer code to adequately cover the input space to a computer code. The strength of Latin Hypercube sampling lies in the ease in which these designs can be generated for any number of factors k with given numbers of levels l; however, very few of these generated designs have desirable properties such as "orthogonality" (Cioppa and Lucas 2007). Orthogonality means that the pairwise correlation between factors is zero. Moreover, the typical application of Latin Hypercubes is to explore quantitative factors to take advantage of the space filling properties inherent in Latin Hypercubes. Recently, researchers at the Naval Postgraduate School have developed the means to generate Latin Hypercube designs for essentially arbitrary numbers of design points for a given number of quantitative, discrete and categorical factors (Naval Postgraduate School 2016) that have, to some approximation, desired properties such as orthogonality and balance.

An early expansion of Latin Hypercube designs were completed by Cioppa and Lucas (2007) who augmented the work of Ye (1998) to produce a larger, though still restricted, catalog of designs constrained to dimension $2^m - 1$ design points by $m + \binom{m-1}{2}$ quantitative factors, where m is integer chosen such that the number of factors $k = 2^m - 2$, and the quantity in parenthesis is the number of combinations taken pairwise. Hernandez et al. (2012), extended the method developed by Cioppa and Lucas via a mixed integer linear program to produce "nearly orthogonal" Latin Hypercubes (NOLH) designs for almost any $k < n$ condition, where n is the number of simulation runs, or design points, in the computing budget. Hernandez et al. (2012) relaxed the requirement for orthogonality and constrained the maximum off diagonal correlation between factors to a small non-zero value; hence, the term "nearly orthogonal" appended to the Latin Hypercube design.

The last extension to this work that we need was made by Vieira et al. (2011; 2013) to produce NOLH mixed designs with good balance and orthogonality. Here, the term "mixed" implies that the factors can be any combination of categorical, discrete or quantitative, and the term "balanced" means that the distribution of factor-levels appear more or less uniformly for a given factor column in the design matrix. The Naval Postgraduate School (2016) has developed an array of Microsoft Excel and Java based tools that automate creation of design matrices based on these expansions (Cioppa and Lucas 2007; Vieira et al. 2011; Hernandez et al. 2012; Vieira et al. 2013) for a nearly arbitrary number of factors. In the following section, we propose a design matrix based on these tools.

4. A Latin Hypercube Design Matrix for NWP

For the design tools created by the Naval Postgraduate School, one tool is the "best" off-the-shelf choice we can make to illustrate our point. Based on the work of Vieira et al. (2013), the "NOB_Mixed_512DP_v1" spreadsheet (Vieira 2012) produces a design for 512 design points, and up to 300 mixed categorical, discrete and continuous factors. It has, however, one limitation for our purpose: for categorical factors, use of this tool is limited to at most 11 levels. Regardless, we will use this tool to produce a design matrix with 512 design points using the 7 categorical factors identified in in Table 1 by restricting the microphysics and planetary boundary layer (PBL) schemes to 11 levels. This restriction is not entirely arbitrary. The WRF-ARW community established a preferential ordering for the namelist options based on suitability criteria that depends on, in part, the domain, the larger scale weather features, etc. For our levels, we will select schemes based on that criteria until we have chosen 11. A portion of the design matrix so created is given as Table 2.

Table 2: First 10 design points of a 512 x 7 design matrix created using NOB_Mixed_512DP_v1.xls (Vieira 2012) for the namelist categories in Table 1.

ra_lw	ra_sw	cu	sf_surface	sf_sfclay	mp	bl_pbl
2	6	2	2	4	5	8
2	2	5	8	2	8	3
2	3	1	1	2	1	8
1	1	4	8	1	11	4
4	4	1	1	3	9	4
6	5	6	6	5	2	3
1	3	3	6	4	6	2
4	1	1	1	7	7	1
6	1	1	2	2	5	4
2	2	1	3	4	3	2

Each row of Table 2 constitutes both a row in our design matrix and a unique φ that we apply to the the mechanism in equation 2 to produce a forecast. For example, the first row means evaluate our mechanism with the long wave radiation scheme set to the second option, the short wave to the sixth option, and so forth. Computing the Pearson correlation coefficient for all 512 design points shows that the design is indeed nearly orthogonal, with a maximum off diagonal correlation coefficient of about 1.25%, and very nearly balanced (each level appears nearly equally often in each column of the design matrix).

Clearly, the approach developed at the Naval Postgraduate School is capable of producing a solution that satisfies our desire to explore the mechanism M. However, the particular tool (Vieira 2012) produces a design with far more design points than we need. Subsequent work will implement the mixed-integer linear programming approach Vieira et al. (2013) developed to directly create a design matrix from the k parameterization classes (θ_k) but with far fewer design points.

5. Summary

We demonstrated that the techniques of experimental design allow us to explore the mechanism for producing a forecast when that mechanism is the WRF-ARW NWP core by showing that we can create a design matrix developed from the various parameterization classes. Thus, we accomplished our main goal with this article. However, these off-the-shelf tools produce designs which contain far more design points than are computationally feasible for our problem. Therefore, to create a design matrix with fewer design points, we must implement the means (Vieira et al. 2013) to construct that matrix. In doing so, we expect to create more compact designs that will also allow us to consider certain combinations of parameterization schemes which the documentation and literature suggests are problematic.

Acknowledgements

We acknowledge the informative discussions with our U.S. Army Research Laboratory colleagues: Dr. Huaqing Cai, Dr. Patrick Haines, Dr. Brian Reen, Dr. Bobby Edmonds, Dr. Robb Randall, Mr. Robert Dumais, Mr. Dave Knapp, and Mr. John Raby. We would also like to acknowledge Dr. Caren Marzban, University of Washington; his exchanges helped clarify our thoughts.

Furthermore, although the workshop has postponed by several months due to weather, the authors would also like to acknowledge the organizers of a Department of Energy/Oak Ridge National Laboratory workshop entitled “Advancing X-cutting Ideas for Computational Climate Science” who accepted an early conceptual version of this work as an “Idea Paper” (Smith and Penc 2015b) for their workshop.

Finally, we would like to acknowledge Dr. Susan Sanchez, Dr. Tom Lucas, and Dr. Andy Hernandez who each have contributed to our understanding of experimental design.

References

Berci, M., V. V. Toropov, R. W. Hewson, and P. H. Gaskell, 2014: Multidisciplinary Multifidelity Optimisation of a Flexible Wing Aerofoil with Reference to a Small UAV. *Structural and Multidisciplinary Optimization*, **50,** 683-699. doi: 10.1007/s00158-014-1066-2.

Box, G. E. P., W. G. Hunter, and J. S. Hunter, 1978: *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* Wiley.

Charney, J. G., 1948: On the Scale of Atmospheric Motions. *Geofysiske Publikasjoner*, **17,** 3-17.

——, 1949: On a Physical Basis for Numerical Prediction of Large-Scale Motions in the Atmosphere. *Journal of Meteorology*, **6,** 371-385.

Cioppa, T. M., and T. W. Lucas, 2007: Efficient Nearly Orthogonal and Space-Filling Latin Hypercubes. *Technometrics*, **49,** 45-55. doi: 10.1198/004017006000000453.

Deming, W. E., 2000: *Out of the Crisis.* MIT Press ed. MIT Press.

Dudhia, J., 2015: Overview of WRF Physics. *2015 Basic WRF Tutorial*, Boulder, CO, National Center for Atmospheric Research.

Fisher, R. A., 1935: *The Design of Experiments.* Oliver and Boyde.

Hall, A., 2014: Projecting Regional Change. *Science*, **346,** 1461-1462. doi: 10.1126/science.aaa0629.

Hernandez, A. S., T. W. Lucas, and M. Carlyle, 2012: Constructing Nearly Orthogonal Latin Hypercubes for Any Nonsaturated Run-Variable Combination. *ACM Transactions on Modeling and Computer Simulation*, **22**. doi: 10.1145/2379810.2379813.

Kleijnen, J. P. C., 2008: *Design and Analysis of Simulation Experiments.* Springer.

Kleijnen, J. P. C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa, 2005: A User's Guide to the Brave New World of Designing Simulation Experiments. *INFORMS Journal on Computing*, **17,** 263-289.

Law, A. M., and W. D. Kelton, 2000: *Simulation Modeling and Analysis.* 3rd ed. McGraw-Hill.

Lorenz, E. N., 1963: Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, **20,** 130-141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Lynch, P., 2006: *The Emergence of Numerical Weather Prediction: Richardson's Dream.* Cambridge University Press.

McKay, M. D., R. J. Beckman, and W. J. Conover, 1979: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, **21,** 239-245. doi: 10.2307/1268522.

Montgomery, D. C., 1997: *Design and Analysis of Experiments.* 4th ed. Wiley.

Naval Postgraduate School, cited 2016: Seed Center for Data Farming. [Available online at http://harvest.nps.edu/.]

Overpeck, J. T., G. A. Meehl, S. Bony, and D. R. Easterling, 2011: Climate Data Challenges in the 21st Century. *Science*, **331,** 700-702. doi: 10.1126/science.1197869.

Pielke, R. A., Sr., 2002: *Mesoscale Meteorological Modeling.* 2nd ed. Academic Press.

Rahimi, A., T. Tavakoli, and S. Zahiri, 2014: Computational Fluid Dynamics (CFD) Modeling of Gaseous Pollutants Dispersion in Low Wind Speed Condition: Isfahan Refinery, a Case Study. *Petroleum Science and Technology*, **32,** 1318-1326. doi: 10.1080/10916466.2011.653701.

Sacks, J., S. B. Schiller, and W. J. Welch, 1989a: Designs for Computer Experiments. *Technometrics*, **31,** 41-47. doi: 10.2307/1270363.

Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn, 1989b: Design and Analysis of Computer Experiments (Includes Comments and Rejoinder). *Statistical Science*, **4,** 409-435. doi: 10.1214/ss/1177012413.

Santner, T. J., B. J. Williams, and W. I. Notz, 2003: *The Design and Analysis of Computer Experiments.* Springer-Verlag.

Schindler, D. E., and R. Hilborn, 2015: Prediction, Precaution, and Policy under Global Change. *Science*, **347,** 953-954. doi: 10.1126/science.1261824.

Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF Version 3. NCAR Technical Note NCAR/TN-475+STR.

Smagorinsky, J., 1963: General Circulation Experiments with the Primitive Equations. *Monthly Weather Review*, **91,** 99-164. doi: 10.1175/1520-0493(1963)091<0099:GCEWTP>2.3.CO;2.

Smith, J. A., and R. S. Penc, 2015a: A Design of Experiments Approach to Evaluating Parameterization Schemes for Numerical Weather Prediction. *Conference on Applied Statistics in Defense*, George Mason University, Fairfax VA, ASA Section on Statistics in Defense and National Security.

Smith, J. A., and R. Penc, 2015b: Using Design of Experiments to Evaluate Numerical Weather Prediction Codes (Accepted Idea Paper). *Advancing X-cutting Ideas for Computational Climate Science*, Rockville, MD, U.S. Department of Energy. doi: 10.13140/RG.2.1.1560.6801.

Stensrud, D. J., 2007: *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models.* Cambridge University Press.

Stevens, B., and S. Bony, 2013: What Are Climate Models Missing? *Science*, **340,** 1053-1054. doi: 10.1126/science.123755.

Vecchi, G. A., and G. Villarini, 2014: Next Season's Hurricanes. *Science*, **343,** 618-619. doi: 10.1126/science.1247759.

Vieira, H., Jr., cited 2016: Nob_Mixed_512dp_V1.xls Design Template. [Available online at http://harvest.nps.edu.]

Vieira, H., Jr., S. Sanchez, K. H. Kienitz, and M. C. Neyra Belderrain, 2011: Generating and Improving Orthogonal Designs by Using Mixed Integer Programming. *European Journal of Operational Research*, **215,** 629-638. doi: 10.1016/j.ejor.2011.07.005.

Vieira, H., Jr., S. M. Sanchez, K. H. Kienitz, and M. C. Neyra Belderrain, 2013: Efficient, Nearly Orthogonal-and-Balanced, Mixed Designs: An Effective Way to Conduct Trade-Off Analyses Via Simulation. *Journal of Simulation*, **7,** 264-275. doi: 10.1057/jos.2013.14.

Warner, T. T., 2011: *Numerical Weather and Climate Prediction.* Cambridge University Press.

Ye, K. Q., 1998: Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments. *Journal of the American Statistical Association*, **93,** 1430-1439. doi: 10.2307/2670057.

Zhu, B., X. Wang, L. Tan, D. Zhou, Y. Zhao, and S. Cao, 2015: Optimization Design of a Reversible Pump-Turbine Runner with High Efficiency and Stability. *Renewable Energy*, **81,** 366-376. doi: 10.1016/j.renene.2015.03.050.