Proceedings of the 1991 Winter Simulation Conference
Barry L. Nelson, W. David Kelton, Gordon M. Clark (eds.)

THE EFFECTS OF BATCHING ON THE POWER OF THE TEST FOR FREQUENCY DOMAIN METHODOLOGY

Douglas J. Morrice

Department of Management Science and Information Systems
The University of Texas at Austin
Austin, Texas 78712-1175

ABSTRACT

This paper considers an implementation of Frequency Domain Methodology (FDM) where input factor levels are held fixed in batches. In addition, the simulation output series is batched and the corresponding batch means are used as observations in the FDM factor screening analysis. For a certain class of metamodels, steady state results are provided. Results are also provided for batch sizes chosen to reduce serial correlation in output series but not necessarily to achieve steady state.

1 INTRODUCTION

With respect to implementation, traditional factor screening techniques and Frequency Domain Methodology (FDM) represent two extremes. For traditional techniques such as a two level factorial design, input factors are *held fixed* during an entire simulation run and *changed across runs.* The mean of the observations from one run is used as one observation in the regression analysis. In FDM, factor levels are *changed during the simulation run* and each output observation generated within a run is used as an observation in the regression analysis.

This paper considers a hybrid implementation whereby input factors levels are fixed in batches (or subruns) of consecutive observations and varied across batches. In addition, the output is batched and the mean values of the output batches are used as observations in the FDM factor screening analysis. Henceforth, we shall refer to this hybrid implementation as *Fixed Factor Batching (FFB).*

Buss (1988) considered batching the output in an FDM without fixing the input factors. He provides favorable results for a single server queueing example. Fixing input factors and batching the output (or making separate simulation runs) was considered by Sanchez (1987) and Sargent and Som (1988). In both papers, factor levels were held fixed in an attempt to achieve steady state within each batch and approximate independence between batches. Both authors rejected FFB in favor of the traditional two level factor screening design because they found the latter to be more efficient when the data are approximately independent and identically distributed (iid). Based on their approach to FFB, we support their conclusions. However, by only batching to achieve steady state and independence between batches, they did not utilize the power of harmonic analysis. In dynamic systems with memory and correlated error, transforming to the frequency domain can 'decouple' terms in the memory filter and diagonalize the covariance matrix of a correlated error process (see Morrice (1990) and Brockwell and Davis (1987), page 130).

The purpose of this paper is to consider FFB with batch sizes chosen to reduce serial correlation and not necessarily to achieve steady state.

2 MODEL ASSUMPTIONS AND BACKGROUND

To illustrate FFB, it is helpful to study a single input factor/single output metamodel of the form:

$$Y(t) = \sum_{r=0}^{q} h(r)X(t-r) + \varepsilon(t). \tag{1}$$

$Y(t)$ is the output, $t = 1, \ldots, N$, is an index which counts the observations generated within a run of the simulation model, N is the total number of observations, $q \leq N$ is the lag length of the memory filter, $h(r)$ is the rth unknown coefficient in the memory filter, $X(t)$ is the input factor, and $\varepsilon(t)$ is a zero mean covariance stationary error process with the following properties:

1. $\varepsilon(t) = \sum_{s=-\infty}^{\infty} w(s)e(t-s)$, where $\sum_{s=-\infty}^{\infty} |w(s)| < \infty$,

2. $\gamma(|i-j|) = Cov(\varepsilon(i), \varepsilon(j))$ is real and $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$,
3. $\{e(t)\}$ are independent with $E(e(t)) = 0$, $\text{Var}(e(t)) = \sigma^2$ and distribution function $\{F_t(e)\}$ where,

$$\sup_{t=1,2,\ldots} \int_{|e|>d} e^2 dF_t(e) \rightarrow 0, \text{ as } d \rightarrow \infty$$

In FDM, the importance of $X(t)$ is tested by making two simulation runs. On the *signal run*, $X(t)$ is varied according to,

$$X(t) = X(0) + a\cos(\omega_1 t) \quad (2)$$

where $X(0)$ is a fixed nominal (or mean) value, a is the oscillation amplitude and $\omega_1 = 2\pi p/N$, for $p \epsilon \{1, \ldots, \lfloor N/2 \rfloor\}$ is the oscillation frequency. Substituting expression (2) into expression (1) yields,

$$\begin{aligned} Y(t) \ = \ & \sum_{r=0}^{q} h(r)X(0) + A(\omega_1)\cos(\omega_1 t) + \\ & B(\omega_1 t)\sin(\omega_1 t) + \varepsilon(t) \end{aligned} \quad (3)$$

where, $A(\omega_1) = \sum_{r=0}^{q} ah(r)cos(\omega_1 r)$ and $B(\omega_1) = \sum_{r=0}^{q} ah(r)sin(\omega_1 r)$. Since the term containing $X(0)$ in expression (3) is of no interest in a factor screening experiment, it will be dropped from the model (only the magnitude of the coefficient of terms that can be altered are of interest in a factor screening experiment). On the *noise (or control) run*, $X(t)$ is fixed at $X(0)$.

For both runs, the *periodogram statistic*, i.e.,

$$I(\omega_1) = (2/N)|\sum_{t=0}^{N-1} Y(t)\exp(-i\omega_1 t)|^2 \quad (4)$$

is computed. (Note: $i^2 = -1$.) The FDM factor screening test is called the *Signal-to-Noise Ratio (SNR)*. It is defined as,

$$SNR(\omega_1) = \frac{I^s(\omega_1)}{I^n(\omega_1)} \quad (5)$$

where the superscripts "s" and "n" denote signal and noise runs, respectively. For large sample sizes, $I^s(\omega_1)$ is approximately distributed as $\alpha_1 \chi^2_{2;\delta_1}$, where

$$\delta_1^2 = \frac{N|G_1|^2}{2\alpha_1}, \quad (6)$$

$$\alpha_1 = \begin{cases} \sum_{|k| \leq \lfloor \frac{N}{2} \rfloor} \gamma(k)\exp(-i\omega_1 k) \\ \sum_{|k| < \frac{N}{2}} \gamma(k)\exp(-i\omega_1 k) + \\ \qquad \gamma(\frac{N}{2})\exp(-i\omega_1(\frac{N}{2})), \end{cases} \quad (7)$$

for N odd and even, respectively and

$$G_1 = a\sum_{r=0}^{q} h(r)\exp(-i\omega_1 r) = A(\omega_1) - iB(\omega_1). \quad (8)$$

$I^n(\omega_1)$ is approximately distributed as $\alpha_1\chi^2_2$ (see Morrice (1990), Section 3.6.2 and Anderson (1971), page 585). Therefore, from expression (5), the SNR has an approximate $F_{2;2,\delta_j}$ distribution. The noncentrality parameter in expression (6) indicates that the FDM hypothesis for testing the importance of $X(t)$ is

$$\begin{aligned} H_O: \ & G_1 = 0 \\ H_A: \ & G_1 \neq 0 \end{aligned} \quad (9)$$

The above analysis can be extended to a multiple input factor/single output metamodel. For large sample sizes, the SNR statistic is approximately independent at distinct frequencies. Therefore, the power of the test for each input factor can be calculated independently of other input factors.

3 A DESCRIPTION OF FFB

The power of the SNR test is a function of N, α_1 and G_1 through δ_1 (see expressions (5) and (6)). In general, batching the output affects the power of the test by replacing N with the number of batches and changing α_1 and G_1; fixing the input factor settings in batches changes G_1. To illustrate, define the regular batch means as: $Z(j) = \{Y((j-1)b+1)+Y((j-1)b+2)+\cdots+Y(jb)\}/b, j = 1, \ldots, m$, where b is the batch size and $m = (N/b)$ is the number of batches. The corresponding batched error process, $\psi(j) = \{\varepsilon((j-1)b+1) + \varepsilon((j-1)b+2)+\cdots+\varepsilon(jb)\}/b$, is a zero mean covariance stationary process (see Law and Kelton (1991), page 555). $\{\psi(j)\}$ also satisfies Properties 1, 2 and 3 given in Section 2 for $\{\varepsilon(t)\}$. Since,

$$\psi(j) = \sum_{s=-\infty}^{\infty} \sum_{k=1}^{b} w(s-k+1)e((j-1)b+1-s)/b \quad (10)$$

and

$$\begin{aligned} & \sum_{s=-\infty}^{\infty} |\sum_{k=1}^{b} w(s-k+1)|/b \\ & \leq \sum_{s=-\infty}^{\infty} \sum_{k=1}^{b} |w(s-k+1)|/b \quad (11) \\ & = b\sum_{s=-\infty}^{\infty} |w(s)|/b < \infty \quad (12) \end{aligned}$$

(Expression 11 follows from the triangle inequality; Expression 12 follows from Property 1 for $\{\varepsilon(t)\}$), therefore Properties 1 and 3 hold for $\{\psi(j)\}$. Property 2 follows because the autocovariance function for $\{\psi(j)\}$ is,

$$\gamma_b(k) = (1/b) \sum_{-(b-1)}^{(b-1)} (1 - |s|/b)\gamma(kb + s) \qquad (13)$$

(see Law and Kelton (1991), page 557).

From expressions (7) and (13), α_1 is modified by batch means to be,

$$\alpha_1^b = \begin{cases} \sum_{|k|\le\lfloor\frac{m}{2}\rfloor} \gamma_b(k)\exp(-i\omega_1 k) \\ \sum_{|k|<\frac{m}{2}} \gamma_b(k)\exp(-i\omega_1 k)+ \\ \qquad \gamma_b(\frac{m}{2})\exp(-i\omega_1(\frac{m}{2})), \end{cases} \qquad (14)$$

for m odd and even, respectively. In the limit,

$$\lim_{b\to\infty}\lim_{m\to\infty} b\alpha_1^b = 2\pi f(0), \qquad (15)$$

where $f(0)$ is the spectral density (see Brockwell and Davis (1987), page 118) of $\{\varepsilon(t)\}$ evaluated at 0 (see Buss (1988), page 551 or Heidelberger and Welch (1981), page 238).

The input factor settings are fixed in batches corresponding to the output batches, and varied across batches according to,

$$X(t) = X(0) + a\cos(\omega_1\lfloor(t + b - 1)/b\rfloor). \qquad (16)$$

For instance, from expression (16), each factor level setting in the set $\{X((j-1)b+s), s = 1, \cdots, b\}$ is equal to $X(0) + acos(\omega_1 j)$, for $j = 1, \cdots, m$. Substituting expression (16) into (1) and computing batch means yields,

$$Z(j) = \sum_{s=0}^{b-1}\sum_{r=0}^{q}(1/b)h(r)\{X(0) + acos(\omega_1(j - \lfloor(r + s)/b\rfloor))\} + \psi(j). \qquad (17)$$

Omitting the terms containing $X(0)$ in expression (17) (see argument following expression (3)) and using the trigonometric identity $cos(\omega(t - r)) = cos(\omega t)cos(\omega r) + sin(\omega t)sin(\omega r)$, expression (17) can be rewritten as,

$$Z(j) = A_b(\omega_1)cos(\omega_1 j) + B_b sin(\omega_1 j) + \psi(j), \qquad (18)$$

where

$$A_b(\omega_1) = \sum_{s=0}^{b-1}\sum_{r=0}^{q}(a/b)h(r)cos(\omega_1\lfloor(r + s)/b\rfloor) \qquad (19)$$

and

$$B_b(\omega_1) = \sum_{s=0}^{b-1}\sum_{r=0}^{q}(a/b)h(r)sin(\omega_1\lfloor(r + s)/b\rfloor) \qquad (20)$$

Since the batch error process satisfies Properties 1, 2 and 3 given in Section 2, then for large m,

$$SNR_b(\omega_1) = \frac{I_b^s(\omega_1)}{I_b^n(\omega_1)} \qquad (21)$$

where

$$I_b(\omega_1) = (2/m)|\sum_{j=0}^{m-1} Z(j)\exp(-i\omega_1 j)|^2 \qquad (22)$$

has an approximate $F_{(2,2;\delta_j^b)}$. The new noncentrality parameter is,

$$(\delta_1^b)^2 = \frac{m|G_1^b|^2}{2\alpha_1^b}, \qquad (23)$$

with,

$$G_1^b = A_b(\omega_1) - iB_b(\omega_1). \qquad (24)$$

G_1^b replaces G_1 in the FDM hypothesis given in expression (9). From expressions (19), (20), and (24),

$$\lim_{b\to\infty} G_1^b = a\sum_{r=0}^{q} h(r). \qquad (25)$$

By fixing factor settings and batching the output, as b approaches infinity, the FDM hypothesis approaches a value which is not a function of a non-zero frequency. For the metamodel in expression (1), the right hand side of expression (25) (without the scaling factor "a") is the coefficient in the metamodel relating the limiting averages,

$$\lim_{N\to\infty} \overline{Y} = \lim_{N\to\infty}\sum_{t=0}^{N} Y(t)/N \qquad (26)$$

and

$$\lim_{N\to\infty} \overline{X} = \lim_{N\to\infty}\sum_{t=0}^{N} X(t)/N. \qquad (27)$$

Traditional factor screening experiments test whether or not this coefficient differs significantly from zero (see Morrice (1990), page 94). Therefore, as b approaches infinity in FFB, the FDM hypothesis approaches the traditional factor screening hypothesis.

4 AN EXAMPLE

Suppose $Y(t)$ and $X(t)$ are related by the following metamodel:

$$Y(t) = X(t) + \theta X(t - 1) + \varepsilon(t) \qquad (28)$$

where,

$$\varepsilon(t) - \phi\varepsilon(t-1) = e(t), \tag{29}$$

θ and ϕ are constants, and $\{e(t)\}$ are iid Normal(0,1).

In this example, power of the test versus frequency curves were constructed for $(\theta, \phi) = (0.5, 0.1), (0.1, 0.5)$, and $(0.9, 0.5)$; these will be referred to as Cases 1, 2, and 3, respectively. In all cases the oscillation amplitude (see Expressions (2) and (16)) was set equal to one, $N = 256$, and batch sizes, $b = 1, 2$, and 4 were considered. An appropriate value for N was chosen using a bound from Morrice (1990) which measures how close the error covariance matrix is to being diagonalized when the analysis is performed in the frequency domain. According to this result, the approximate distributional assumptions for the SNR_b test (see Expression (21)) are valid for $m = 256, 128$ and 64. The power of the test curves were generated using the central F approximation of the non-central F found in Scheffe (1959) on page 144. The FORTRAN IMSL routine FDF (see IMSL/STAT LIBRARY, page 925) produced the values for the cumulative distribution function of an F-distribution.

Figures 1, 2, and 3 contain the power of the test curves for Cases 1, 2, and 3, respectively. In each of the diagrams the label "bsiz" is the batch size, b. As expected, in all three cases, batching tends to flatten out the power of the test curves. This follows because both numerator and denominator of the noncentrality parameter are approaching constant values as the batch size increases (see expressions (15) and (25)). The shapes of the power of the test curves are determined by the interaction of α_1^b and $|G_1^b|^2$ in the noncentrality parameter δ_1^b.

For Case 1 (see Figure 1), since ϕ is small, $\{\varepsilon(t)\}$ is close to an iid Normal(0,1) process. Batching to produce $\{\psi(j)\}$ further reduces the correlation. Since an uncorrelated process has a flat spectral density (see Brockwell and Davis (1987), page 121), α_1^b is close to being constant across all frequencies for all three batch sizes. The shape of the power of the test curves in Figure 1 are, for the most part, determined by the shape of $|G_1^b|^2$. For $b = 1, |G_1^b|^2$ is a scaled MA(1) spectral density which has a maximum value for frequency zero and decreases as the frequency increases over the interval (0,0.5] (see Brockwell and Davis (1987), page 122). As b increases, $|G_1^b|^2$ approaches this maximum value across all frequencies (see expression (25)). Therefore, as b increases, the power of the test improves across all frequencies.

Case 2 is the opposite of Case 1 (see Figure 2). Since θ is small, $|G_1^b|^2$ is close to being constant across all frequencies for $b = 1, 2$, and 4. For $b = 1, \alpha_1^b$ is approximately equal to a scaled spectral density for an AR(1) process. This function has a maximum at zero and decreases over the frequency range (0,0.5]. When b increases (and the corresponding value of m is large enough), α_1^b approaches this maximum value across all frequencies (see expression (15)). Since α_1^b appears in the denominator of δ_1^b, the power of the test curve decreases as α_1^b increases. In this case, increasing b decreases the power of the test across all frequencies.

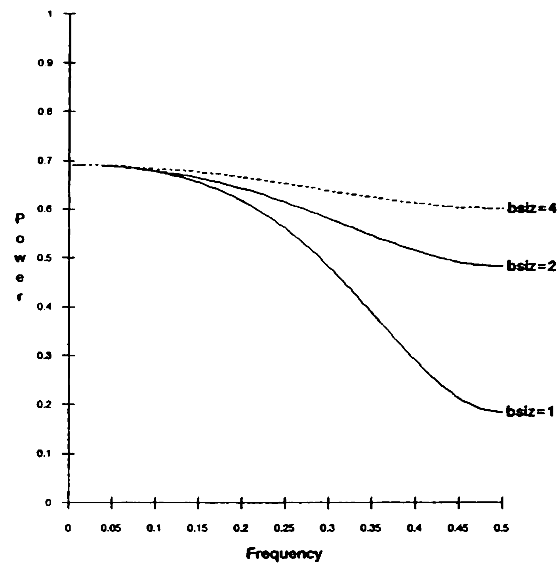

Figure 1: Test Power Versus Frequency For Case 1

Case 3 illustrates that increasing the value of b may not uniformly increase or decrease the power of the SNR_b test. (compare power of the test curves for batch sizes one and two in Figure 3). It also illustrates that choosing an intermediate value for b may be more desirable than batching to achieve steady state (Case 1) or not batching at all (Case 2). In Figure 3, unlike the power of the test curve for $b = 1$, the power of the test curve for $b = 2$ does not have a region where the power of the test is close to zero (compare these power curves for frequencies higher than 0.4). In addition, the power of the test curve for $b = 2$ almost uniformly dominates the curve for $b = 4$.

5 CONCLUSION

In this paper, we have examined some of the properties of FFB and FDM analysis. For classes of metamodels discussed in Section 2, steady state results are provided. In Section 4, an example illustrates cases where FFB improves the power of the FDM hypothesis test and cases where it does not. For cases where FFB is beneficial with respect to the power of the test, we have illustrated that batching to an inter-

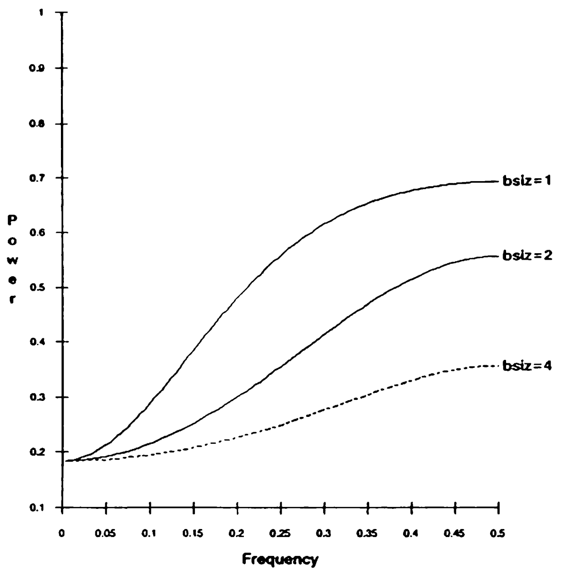

Figure 2: Test Power Versus Frequency For Case 2

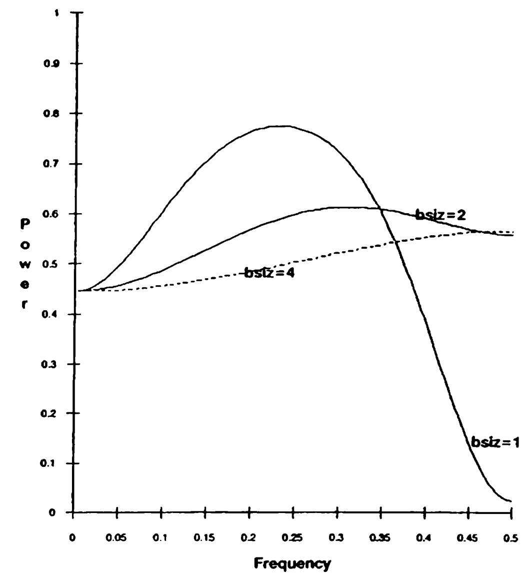

Figure 3: Test Power Versus Frequency For Case 3

mediate batch size may be better than batching to achieve steady state.

Future research includes considering ways to use the structure of a simulation model to determine if FFB improves FDM. For example, information such as whether or not an input factor signal must pass through a waiting line may provide insights into the usefulness of FFB. Another research topic concerns the effect of batch means on minimum required run length for FDM. A measure for minimum required run length, given by Morrice (1990), depends on the autocovariance function of the error process which is changed by batching. Finally, based on the results of Buss (1988), in future research we will also consider the effects of chosing a batch size for the input factor settings which is different from the output batch size.

ACKNOWLEDGEMENTS

The author thanks Lee Schruben for his many helpful comments.

REFERENCES

Brockwell, P.J. and Davis, R.A. 1987. *Time Series: Theory and Methods.* New York: Springer-Verlag.

Buss, A.H. 1988. Some extensions and limitations of frequency domain experiments. In *Proceedings of the 1988 Winter Simulation Conference,* eds. M. Abrams, P. Haigh and J. Comfort, 549-557. Institute of Electrical and Electronics Engineers, San Diego, California.

Heidelberger, P. and Welch, P. 1981. A Spectral Method for Confidence Interval Generation and Run Length Control in Simulations. *Communications of the Association for Computing Machinery* **24**: 233-245.

IMSL 1987. *IMSL STAT/LIBRARY: FORTRAN Routines for Statistical Analysis (Version 1.0).* Distributed by International Mathematical and Statistical Libraries, Inc., Houston, Texas.

Law, A.M. and Kelton, W.D. 1991. *Simulation Modeling and Analysis,* Second Edition. New York: McGraw-Hill.

Morrice, D.J. 1990. Simulation Factor Screening. Ph.D. Dissertation, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

Sargent, R.G. and Som, T.K. 1988. A Different View of Frequency Domain Experiments. Working Paper #88-01, Syracuse University, Syracuse, New York.

Scheffe, H. 1959. *The Analysis of Variance.* New York: John Wiley and Sons.

AUTHOR BIOGRAPHY

DOUGLAS J. MORRICE is an assistant professor in the Department of Management Science and Information Systems at The University of Texas at Austin. His research interests are in the statistical design and analysis of large scale simulation experiments and the statistical aspects of quality control. He is a member of the The Institute of Management Science and the Operations Research Society of America.