Power spectral artifacts in published balloon data and implications regarding saturated gravity wave theories

E. M. Dewan
Air Force Research Laboratory, Space Vehicles Directorate, Hanscom Air Force Base, Massachusetts

N. Grossbard
Boston College, Institute for Scientific Research, Chestnut Hill, Massachusetts

Abstract. There are several theories of atmospheric gravity wave power spectral densities (PSDs) which have been published. These, in turn, have inspired numerous experimental tests. The spectra involved are in the class denoted “stochastic, red noise spectra.” This means that most of the power is at the low-frequency end and obeys a power law falloff in going to higher frequency. The present paper describes how some published experimental spectra are flawed by an artifact of spectral analysis which has not heretofore been recognized in the literature. It involves both an amplitude fluctuation enhancement and a coupling between spectral amplitude and slope, and it can be avoided only by stringent control of spectral leakage. Because of “trade-off” considerations every data set, depending on its length and signal characteristics, requires a different method of analysis. It is therefore required that PSD analysis programs must be adjusted and tested to fit each situation. For this purpose a simple method is described to simulate data of known general characteristics for test purposes (to avoid the pitfalls). Since the papers by Nastrom et al. [1997] and de la Torre et al. [1997] have the unfortunate artifact in their analyses, their conclusions regarding saturated gravity wave theories should be reexamined.

1. Introduction

The discovery [Dewan et al., 1984] that the horizontal velocity vertical wavenumber spectrum appears to be “universal” has inspired a number of theories to be formulated in order to explain this finding. These include those of Dewan and Good [1986], extended by Smith et al. [1987], as well as Weinstock [1990], Hines [1991a, b, c], Gardner [1994], and Medvedev and Klaassen [1995]. This discovery was predicted by Van Zandt [1982] in a seminal paper based, in part, on the empirical model of Garrett and Munk [1975] for internal ocean waves. More recently, numerous observations have been made, some of which support the universality idea and some of which do not. For example, a recent paper by Nastrom et al. [1997] has raised various questions about saturation theories of gravity waves. In particular, they showed “scatter diagrams” that display correlations between spectral slope and spectral amplitude. As we shall show, their results are contaminated by an artifact not previously discussed in the literature; and this was the original motivation of the present investigation.

In recent years, many “canned programs” and “recipes” have been published which perform spectral analysis of experimental data. Unfortunately, such publications give little hint of the vast number of pitfalls that are involved, and they do not give much information about the literature available which might be helpful in this regard. On the basis of our own experience and that of our colleagues, we are very much aware that there is no single procedure (due to trade-offs) which is appropriate to every data set, and that therefore in our joint opinion it would be very beneficial if every procedure were tested with data of known spectral characteristics before publication. Naturally, the simulated data must have the same number of points and general characteristics of the data to be analyzed. If this test procedure is omitted, one places oneself in a position similar to a concert violinist who does not bother to tune his instrument before giving a public performance. Anyone who tests his spectral analysis programs knows very well what is behind this statement. The nature of the data set and the goal of the researcher are crucial for the choice of method of spectral analysis [e.g., Percival and Walden, 1993].

How are such simulations to be carried out? In the case at hand, where we are concerned with “red noise” spectra, the answer is easy to state: “Just filter white noise from a Gaussian white noise generator with the filter that corresponds to the desired spectrum.” This sentence does not answer the questions of how to actually do this in a practical and simple manner. Section 2 below answers this question in detail. It is the technique used in the remainder of our study.

Section 3 describes a method of spectral analysis that as we found, eliminated the artifact mentioned above. It involves “prewhitening” by means of “first-differencing.” This, of course, would not be guaranteed to work for other classes of data. For example, the present case involves spectra not much steeper than a slope of −3. If they were significantly steeper, something more powerful than prewhitening by first differencing would be required. Percival and Walden [1993] explain more general and powerful methods of prewhitening for such occasions.

Section 4 shows our simulations with and without leakage reduction and exhibits the new artifact, while section 5 examines these results more quantitatively. The goal of the analysis
The problem by performing the simulation with a set of 16,384
individual sine waves with suitably randomized amplitudes and phases
instead, we recommend the following simpler approach:

The first step is to simulate a pseudo-random Gaussian
white noise data stream using standard techniques (see Press
[1994] or, alternatively, Knuth [1981]). Having obtained 16,384
values, we used a discrete Fourier transform (DFT) (using the
fast Fourier transform method) to obtain 16,384 complex Fourier
amplitudes in wavenumber space (with positive and negative
wavenumbers). These amplitudes were subsequently multiplied
by factors so that the PSDs would, on average, have the
spectrum. These amplitudes were subsequently multiplied
by the fast Fourier transform method) to obtain 16,384 complex
Fourier amplitudes in wavenumber space (with positive and neg-
ative wavenumbers). These amplitudes were subsequently multi-
plied by factors so that the PSDs would, on average, have the

where the superscript on $\Psi$ is to remind one that it is a two-
sided PSD.

Let $A(k_n)$ be the complex amplitudes of the Fourier-
transformed white noise. These will be multiplied by coefficients
$B(k_n)$, so

$$\Psi^{(2)}(k_n) = \Psi^{(2)}_{H}(k_n),$$

on average, where $\Psi^{(2)}_{H}(k_n)$ is the “model” or desired spec-
trum. Thus the first thing that must be done is to obtain the values
of $B(k_n)$ which will accomplish this purpose.

First we must associate the $A(k_n)$ amplitudes (from white
noise) to its PSD, which we will call $\Psi^{(2)}_{H}(k_n)$. We start by
considering a single Fourier component at frequency $k_n =
(n'/N\Delta x)$. (Let $x_j = j\Delta x$.)

$$f(x_j) = A(k_n)e^{i(2\pi/n')(j\Delta x)/(N\Delta x)},$$

Inserting (3) into (1),

$$\Psi^{(2)}_{H}(k_n) = \Delta x N \sum_{j=0}^{N-1} A(k_n)e^{i(2\pi/n')(j\Delta x)/(N\Delta x)}^2,$$

where the subindex $n'$ on the left reminds us that this PSD is for
a single Fourier component with wavenumber $k_n' = n'/N\Delta x$.

This can be written

$$\Psi^{(2)}_{H}(k_n) = \frac{\Delta x}{N} A^2(k_n) \delta(n - n') N^2.$$

(Take to see that (5) is correct, see any detailed discussion of the
DFT such as found in the works of Percival and Walden [1993,

Thus

$$\Psi^{(2)}_{H}(k_n) = A^2(k_n) N\Delta x.$$  

(6)

The next step is to obtain the two-sided PSD of the band-
limited ($-k_{max}$ to $+k_{max}$) Gaussian white noise, which is our
starting set of simulated data. Let its variance be called $\sigma^2$. The
reader will recall that in the continuous case,

$$\sigma^2 = \int_{-k_{max}}^{k_{max}} \Psi^{(2)}_{H}(k) dk,$$

which is to say that the total variance of the noise would equal
the integral of the two-sided “variance density spectrum” over the
entire negative and positive wavenumber range of our band-
limited noise. Translating this to our discrete case,

$$\sigma^2 = \sum_{n=-\lfloor N/2 \rfloor + 1}^{\lfloor N/2 \rfloor} \Psi^{(2)}_{H}(k_n) \Delta k,$$

where $\Delta k = 1/N\Delta x$ and where we recall that the maximum
positive wavenumber (or Nyquist frequency) is labeled $n =
N/2$. 

of section 5 is to display the quantitative nature of the new
artifact within the context of gravity wave data and their spectra.
The range of the slopes considered are in the range of $-2.5$
to $-3.5$, which should cover all relevant slopes since those that
are less steep would probably not show the artifact. It should
be mentioned that spectra with positive slopes are not consid-
ered here. Obviously, averages rather than differences, i.e.,
low-pass rather than high-pass methods, would be needed for
the latter. The reader who wishes to explore the artifact further
for other types of spectra can use the simulation technique
described below to do so. Our immediate and more pressing
purpose is, of course, to recognize the problem and, more
importantly, to eliminate it when it arises. In the conclusion
(section 6) we mention the fact that Nastrom et al. [1997] must
repeat their analysis with due attention to spectral leakage
before their own conclusions can be accepted (especially in
regard to their slope-log amplitude “scatterplots”).

2. Numerical Simulations of Stochastic Data
Possessing Known Power Spectral Density (PSD)
Characteristics

Previously, descriptions [Owens, 1978; Fairall and White,
1991] have been published which explain how to simulate
“data” with known PSD shapes. These methods failed to take
into account the fact that no resolvable sinusoidal components
are permitted if a true stochastic simulation is desired. This
is due to the fact that a periodogram of a resolvable sine wave
component will double its amplitude if the database is doubled
in length, whereas doubling the length of a sample of noise will
not have such an effect. [See Gottman, 1981, p. 241.]

Then there is a second problem usually not mentioned in
papers on simulating “data” with known PSD shapes. The
simulation of a finite set of data introduces a bias [Oppenheim
and Schafer, 1975, p. 543, equation (11.33)]. The bias is caused
by the fact that the PSD of a finite data set is a distorted
version of the intended form due to the fact that it involves a
convolution between the theoretical spectrum and the digital
sine function. The longer the data stream the more the sine
effectively becomes a delta function. For this reason we simu-
late data sets much longer than the subset we analyze. In what
follows we will use a method of simulation which avoids this
problem by performing the simulation with a set of 16,384
points and then analyzing only a subset of 128 data points.
(The number 128 was chosen because it is very close to the size
of some published balloon data reports).

It would be inconveniently difficult to simulate 8192 individ-
ual sine waves with suitably randomized amplitudes and phases
with known PSD characteristics if one were to do so on an
individual sine wave basis. Instead, we recommend the follow-
ing simpler approach:

The first step is to simulate a pseudo-random Gaussian
white noise data stream using standard techniques (see Press
[1994] or, alternatively, Knuth [1981]). Having obtained 16,384
values, we used a discrete Fourier transform (DFT) (using the
fast Fourier transform method) to obtain 16,384 complex Fourier
amplitudes in wavenumber space (with positive and negative
wavenumbers). These amplitudes were subsequently multiplied
by factors so that the PSDs would, on average, have the
desired shape. We shall call $\Psi_{H}(k_n)$ the “model spectrum” we
wish to simulate.

Now let us consider the mathematical details. In the follow-
ing we will use the two-sided PSD for mathematical analysis
and the one-sided PSD for presentation of results. The data set
to be simulated will be a function of discrete position $x_j$. Let
this $x_j$ be given at spacing $\Delta x$ and let $j$ take on $N$ values, $j =
0, 1, \ldots , N-1$. The data set will be called $f(x_j)$. The discrete
two-sided PSD of $f(x_j)$ at wavenumber $k_n = (n/N\Delta x)$ is given by

$$\Psi^{(2)}(k_n) = \frac{\Delta x}{N} \sum_{j=0}^{N-1} f(x_j) \exp \left[ -\frac{2\pi i n \cdot j}{N} \right]^2,$$  

(1)

where

$$\Psi^{(2)}_{H}(k_n) = \Psi^{(2)}_{H}(k_n),$$

on average, where $\Psi^{(2)}_{H}(k_n)$ is the “model” or desired spec-
trum. Thus the first thing that must be done is to obtain the values
of $B(k_n)$ which will accomplish this purpose.

First we must associate the $A(k_n)$ amplitudes (from white
noise) to its PSD, which we will call $\Psi^{(2)}_{H}(k_n)$. We start by
considering a single Fourier component at frequency $k_n =
(n'/N\Delta x)$. (Let $x_j = j\Delta x$.)

$$f(x_j) = A(k_n)e^{i(2\pi/n')(j\Delta x)/(N\Delta x)},$$

Inserting (3) into (1),

$$\Psi^{(2)}_{H}(k_n) = \Delta x N \sum_{j=0}^{N-1} A(k_n)e^{i(2\pi/n')(j\Delta x)/(N\Delta x)}^2,$$

where the subindex $n'$ on the left reminds us that this PSD is for
a single Fourier component with wavenumber $k_n' = n'/N\Delta x$.

This can be written

$$\Psi^{(2)}_{H}(k_n) = \frac{\Delta x}{N} A^2(k_n) \delta(n - n') N^2.$$  

(5)

To see that (5) is correct, see any detailed discussion of the
DFT such as found in the works of Percival and Walden [1993,

Thus

$$\Psi^{(2)}_{H}(k_n) = A^2(k_n) N\Delta x.$$  

(6)

The next step is to obtain the two-sided PSD of the band-
limited ($-k_{max}$ to $+k_{max}$) Gaussian white noise, which is our
starting set of simulated data. Let its variance be called $\sigma^2$. The
reader will recall that in the continuous case,

$$\sigma^2 = \int_{-k_{max}}^{k_{max}} \Psi^{(2)}_{H}(k) dk,$$

which is to say that the total variance of the noise would equal
the integral of the two-sided “variance density spectrum” over the
entire negative and positive wavenumber range of our band-
limited noise. Translating this to our discrete case,

$$\sigma^2 = \sum_{n=-\lfloor N/2 \rfloor + 1}^{\lfloor N/2 \rfloor} \Psi^{(2)}_{H}(k_n) \Delta k,$$

where $\Delta k = 1/N\Delta x$ and where we recall that the maximum
positive wavenumber (or Nyquist frequency) is labeled $n =
N/2$.
Since $\Psi_{\text{noise}}^{(2)}(k_n)$ is white (i.e., nominally a constant, though in reality varying about the constant), one can write

$$\sigma^2 = \Psi_{\text{noise}}^{(2)}(k_n) \cdot \frac{1}{N\Delta x} \cdot N$$

or

$$\sigma^2 = \frac{\Psi_{\text{noise}}^{(2)}(k_n)}{\Delta x},$$

hence

$$\Psi_{\text{noise}}^{(2)}(k_n) = \sigma^2(\Delta x) \left( \text{for } n = -\frac{N}{2} + 1 \text{ to } \frac{N}{2} \right).$$

This is the PSD of our original Gaussian band-limited white noise simulated data set. We now turn to the process of “molding” it or “filtering” it in wavenumber space.

We start by using (6), which relates each $A(k_n)$ component to a PSD, $\Psi_{\text{noise}}^{(2)}(k_n)$, and also uses the fact that our $A(k_n)$ values were created by noise and hence have the PSD given by (11). Thus one obtains

$$A^2(k_n) N\Delta x = \Delta x \sigma^2,$$

or solving for $A^2(k_n)$,

$$A^2(k_n) = \frac{\sigma^2}{N} \left( \text{for } n = -\frac{N}{2} + 1 \text{ to } \frac{N}{2} \text{ or } n = 0 \text{ to } N - 1 \right).$$

To shape the values of $A(k_n)$ so that the associated PSD will have the shape of $\Psi_{\text{noise}}^{(2)}(k_n)$, the model spectrum, we multiply each complex $A(k_n)$ by a factor $B(k_n)$ and do this for all $k_n$ values. The question now to answer is “what values does one use for $B(k_n)$?”

From (6) we have for general $k_n$,

$$A^2(k_n) = \frac{\Psi_{\text{noise}}^{(2)}(k_n)}{(N\Delta x)}. \quad (14)$$

Thus to create data with PSD equal to $\Psi_{\text{noise}}^{(2)}(k_n)$, one must multiply the $A(k_n)$ values by $B(k_n)$ values such that

$$B^2(k_n) A^2(k_n) = \frac{\Psi_{\text{noise}}^{(2)}(k_n)}{N\Delta x}. \quad (15)$$

Solving for $B^2(k_n)$,

$$B(k_n) = \sqrt{\frac{\Psi_{\text{noise}}^{(2)}(k_n)}{N\Delta x A^2(k_n)}}, \quad (16)$$

using (13) for $A^2(k_n)$ in (16),

$$B(k_n) = \frac{\Psi_{\text{noise}}^{(2)}(k_n)}{\Delta x (\sigma^2)}.$$

In the next section we will be using the one-sided model spectrum:

$$\Psi_{\text{model}}^{(2)}(k_n) = 2\Psi_{\text{noise}}^{(2)}(k_n) \quad (\text{for } n = 0 \text{ to } \frac{N}{2}, \text{i.e., } k_n \text{ positive definite}),$$

ior solving for $A^2(k_n)$,

or solving for $A^2(k_n)$,

or solving for $A^2(k_n)$,

or solving for $A^2(k_n)$,

or solving for $A^2(k_n)$,

or solving for $A^2(k_n)$,
Thus to obtain the desired simulated data, one performs an inverse DFT on the \( A(k_n) B(k_n) \) products. Let \( g(x) \) represent this simulated data with PSD \( C(k) \). Then

\[
g(x_i) = \frac{1}{N\Delta x} \sum_{n=0}^{N-1} A(k_n) B(k_n) e^{2\pi i n \ell / N}. \tag{20}
\]

In our simulations, \( \ell = 0, 1, \cdots, 16,384 \). This gives two advantages: (1) none of the Fourier components will be resolvable in the analysis of only 128 of these points of \( g(x) \) (or the other cases described below) and (2) the large number of frequencies and long database allow us to create simulated data with as faithful a representation of the desired spectrum as possible. (We subsequently learned that far fewer points may actually be needed. That is to say, our experiments suggest that if \( N \) is the number of points, then a simulation with a small number times \( N \) points would probably be accurate enough.)

### 3. Power Spectrum Analysis Procedure

The following method to arrive at spectral estimates of data, when the prewhitening and postdarkening steps are included, is very effective and appropriate for the case under discussion, i.e., small samples (128 points) of red noise. However, in the case of the published balloon data mentioned earlier, the prewhitening and postdarkening steps were omitted. In what follows we shall contrast the two procedures when applied to our simulations.

#### 3.1. Trend Removal

A least squares fitted line to the data was subtracted from the data. In this way we were able to obtain a time series that was stationary.

#### 3.2. Prewhitening

As was mentioned, this step will be left out, and then we will compare the results to the cases where it is incorporated. In those cases where it is left out, the “postdarkening” step is also omitted [see Blackman and Tukey, 1959]. There are many ways to prewhiten data; but with the data under consideration, we found that “first differencing” was adequate for our purpose. Thus if \( x_i \) represents the data after trend removal, then \( P_i \), the prewhitened version, is given by

\[
P_i = x_{i+1} - x_i, \tag{21}
\]

for \( i = 1, \cdots, N - 1, P_N = 0 \).

#### 3.3. Discrete Fourier Transform (DFT)

The Fourier transform of the prewhitened data, \( F(k) \), is

\[
F(k_n) = \sum_{j=0}^{N-1} P_i \exp \left( -2\pi i j k_n / N \right), \tag{22}
\]

where as before, \( k_n = (n/N\Delta x) \).

#### 3.4. Raw Periodogram

The raw periodogram was obtained from

\[
\Psi^{(1)}(k_n) = \frac{2\Delta x}{N} |F(k_n)|^2. \tag{23}
\]

Note that this is the one-sided version of the spectrum.

#### 3.5. Hanning

The PSD was smoothed (for the purpose of reducing variability) by using a weighted average over neighboring frequencies, giving us the so-called Hanned PSD, \( \Psi^{(1)}(k_n) \).

\[
\Psi^{(1)}(k_n) = \frac{1}{4} \Psi^{(1)}(k_{n-1}) + \frac{1}{2} \Psi^{(1)}(k_n) + \frac{1}{4} \Psi^{(1)}(k_{n+1}) \tag{24}
\]

for \( k_1 \leq k_n \leq k_{N/2-1} \).

#### 3.6. Postdarkening

To compensate for the “first-difference” form of prewhitening, when it was employed, we used

\[
\Psi^{(1)}_{pd}(k_n) = \frac{1}{2(1 - \cos (2\pi k_n \Delta x))} \tag{25}
\]

for \( k_1 \leq k_n \leq k_{N/2} \).
4. Comparisons of Slopes and Amplitudes and the Discovery of the Anomalous Fluctuation Effect

4.1. Input PSDs

Figure 1 shows an example of simulated data using the following input PSDs and the above procedures. This used the input model spectrum given by

\[ C_M \sim 1 \cdot k^5 N^2 k_n^2 \sim k_{\text{max}}^3, \tag{26} \]

where \( N^2 = 0.02 \text{ rad/m}, k_* = 2.03 \times 10^{-3} \text{ rad/m} \) (i.e., \( \lambda_* = 3 \text{ km}, \) and \( k_* = 2\pi/\lambda_* \) where \( \lambda_* = N\Delta x/n). \) In these simulations we considered mainly two cases in regard to the power below \( k_* \). Case 1 is where the power was set equal to the value it had at \( k_* \) for all \( k < k_* \) to the lowest simulated frequency. Case 2 is where \( \Psi_M(k_n) = 0 \) for all \( k < k_* \) to the lowest simulated frequency. A third case is also mentioned below. The value of \( \sigma^2 = 1 \) for the pseudo-random white noise was used for these simulations.

Figure 2 shows an example of a PSD that has been fitted with a line on the log-log plot by least squares, over the wavenumber range of \( 6.85 \times 10^{-3} \text{ to } 3.14 \times 10^{-2} \text{ rad/m}. \) The data spacing was \( \Delta x = 50 \text{ m}, \) thus the Nyquist wavenumber was \( 2\pi/100 \text{ m} = 6.28 \times 10^{-2} \text{ rad/m}. \) In Figure 2 the fitted line is dashed. From the fitted line a representative PSD amplitude was ascertained. To avoid amplitude swings due mainly to slope variation, we used the appropriate “center of mass” wavenumber:

\[ \log_{10} k_{\text{CM}} = \left( \frac{1}{\text{number of } k_n \text{ in the fitted range}} \right) \cdot \sum_{k_n \text{ in the fitted range}} \log_{10} (k_n). \tag{27} \]

The “center of mass” wavenumber \( k_{n\text{CM}} \) is therefore

\[ k_{n\text{CM}} = 10^{\log_{10} k_{\text{CM}}}. \tag{28} \]

Thus we found \( k_{n\text{CM}} = 1.76 \times 10^{-2} \text{ rad/m}. \)

Figure 3 shows a scatterplot, giving \( \log \Psi \) at \( k_{n\text{CM}} \) versus the slope of the fitted line. The values of \( \log \Psi (k_{n\text{CM}}) \) are taken from the fitted lines. In all, there are 162 simulated spectra represented in this figure.

4.2. Study of Scatterplots

Each spectrum in Figure 3 (and all subsequent plots if not otherwise indicated) was based on the first 128 points of the simulated data. When prewhitening was omitted, we found that the large amount of variability to be seen in this figure is larger than expected, considering the fact the data being simulated have a unique amplitude and slope. To explore this variability further, we changed the model (26) in one respect. We changed the exponent of \( k_n \) from 3 to 3.5. Nothing else was changed. The resulting scatterplot is given in Figure 4, and as can be seen from the comparison between Figures 4 and 3, the variability in Figure 4 is dramatically larger. In Figure 5, a \(-2.5\) exponent was used, and the variability dropped equally dramatically.

There is a well-known theory to explain spectral variability. It is to be found in such books as Jenkins and Watts [1968, pp.
Percival and Walden [1993, pp. 220–228], and Otnes and Enochson [1992, pp. 206–223]. This theory assumes that the data approximate white Gaussian noise that is stationary. Under such assumptions it can be shown that for the case of a raw periodogram the distribution of variance is the same as a “chi squared” distribution with two degrees of freedom. It has also been shown that for the general case (where such things as smoothing the PSD, or averaging several spectra together in order to decrease variance have been carried out), one has the ratio

\[
\frac{\text{PSD measured}}{\text{PSD theoretical}} \approx \chi^2_n \nu, \quad (29)
\]

where \(\approx\) means “has the same distribution” and where \(\nu\) is the number of degrees of freedom. (The more smoothing and averaging are involved, the higher the \(\nu\) becomes.)

The analytic form for \(\chi^2/\nu\) is given by

\[
\left(\frac{\chi^2}{\nu}\right) = \left(\frac{\nu}{2^\nu (\nu - 1)!}\right) \nu \left(\nu^2 (\nu - 1)!\right)^{x/2} e^{-\nu x/2}. \quad (30)
\]

Equation (30) is used below for our plots. Since Hanning was used in our smoothing procedure, we will plot the above for \(\nu = 6\), which is the approximate number of degrees of freedom for Hanning.

Figure 6 shows the distribution of (PSD measured/PSD theoretical) for the \(-3\) slope simulations of Figure 3. The PSD measured value is now taken from the PSD itself rather than

Figure 4. Scatterplot similar to Figure 3 but with input spectral slope of \((-3.5)\). Note increase of slope of fitted line here. (Again, no prewhitening.)

Figure 5. Scatterplot as in Figure 4 but with input spectral slope of \((-2.5)\).
from the fitted line. The dotted line is the plot for \((x_6^2)/6\) (i.e., (30) with \(r = 6\)). Notice that the theory would predict that the two curves would have roughly equal values, whereas the experimental variability far exceeds theoretical prediction. The problem with the theory lies in its assumption of white noise, whereas our PSD represents steep red noise.

Figures 7 and 8 show plots similar to Figure 6 for the \(-3.5\) and \(-2.5\) slopes (simulated in Figures 4 and 5, respectively). The disagreement between the \(\chi^2\) distribution and the measured distribution increased with an increase of steepness of slope and vice versa; and this dependence on slope appears to be quite sensitive. This anomalous variability of the PSD (with respect to standard theory) seems to have been generally overlooked until now. The underlying cause of the increase of variability with increased steepness of spectral slope is directly linked to the leakage effect.

While the above procedure for spectral analysis (i.e., without prewhitening) involved “windowing,” it did not offer much protection against leakage, and this shows up in the average spectral slopes measured. A \(-3.5\) slope simulation resulted in an average measured slope of \(-2.88\), a \(-3.00\) input gave a measured \(-2.72\) slope, and \(-2.5\) input gave \(-2.40\) (the effect diminished greatly with the steepness). This further indicates that this method of spectral analysis is deeply flawed when used on steep spectra and that the flaw is due to leakage.

Traditionally, the best defense against leakage is “prewhitening” [see Blackman and Tukey, 1959; Percival and Walden, 1993]. Briefly, this technique attempts to make the spectrum as constant or white as possible. For example, if one were to cause a spectrum with a \(-3.0\) slope to be analyzed as a zero sloped spectrum, then there would be total and perfect control over leakage. In the prewhitening procedure described by (21), a spectrum with slope \((-n)\) would, approximately, become a spectrum with slope \(-n^2\), thus a \(-3\) sloped spectrum would become \(-1\), which while not perfect seems to be sufficient for present purposes. As has already been mentioned, there are more involved methods of prewhitening, which can do virtually any degree of prewhitening needed, and they are described by Blackman and Tukey [1959] and Percival and Walden [1993].

When prewhitening (and postdarkening) are performed (equations (21) and (25)), the anomalous spectral variability vanishes. See Figures 9, 10, and 11, which give scatterplots and \(\chi^2\) plots for input slopes of \(-3.5\), \(-3.0\), and \(-2.5\), respectively. Not only do the \(\chi^2\) distributions match the measured distributions of (PSD measured/PSD actual) but also the average slopes match the input (simulation) slopes (e.g., \(-3.46\) measured versus \(-3.5\) input, \(-2.95\) for \(-3.0\) and \(-2.45\) for \(-2.5\)). Of course, the prewhitened data, when spectrally analyzed, now fit more closely the assumptions of the theory of spectral variability (e.g., the white noise assumption), and it should therefore come as no surprise that this cures the variability problem.

Figure 12 shows the scatterplot and \(\chi^2\) plot for the case...
where the input slope was $-3.0$, and where a straight line connecting $0$ to $k_+$ completes the model of (26). No prewhitening was used here. As can be seen, the scatterplot has a fitted slope of about 0.27. The spread and slope of the line is comparable to some balloon data of Nastrom et al. [1997]. Clearly, the line and its slope for the scatterplot would have no statistical significance in the present instance since only one input slope and amplitude is involved. This would raise serious questions about any conclusions one could arrive at from experimental data with samples having a low number of points (128), which looked like Figure 12, and was arrived at without prewhitening. As can be seen, the anomalous variability can be very misleading indeed.

In the high-resolution spectral analysis of smoke trail data [Dewan et al., 1984, 1988] the number of points per profile was of the order of 1000. More precisely, there were 2–5 segments containing about 400 points each. Figure 13 shows a simulation identical to that in Figure 12 but now with 1000 points. Note the tremendous decrease in the variability observed both in amplitude and slope. Figure 14 shows what happens when prewhitening is included, as was done by Dewan et al. [1984, 1988]. In this case, the slope of the fitted line in the scatterplot is essentially zero. (In the work of Dewan et al. [1988] the details of a PSD analysis procedure is given, which for this type of data seems to be accurate and robust against artifacts.)

5. Quantitative Observations

In the following we show some quantitative observations regarding the reduction of the fluctuation artifact by means of (1) prewhitening, (2) increasing the number of data points, and (3) data tapering. These techniques are all applied to data simulations.

First, let us consider the effects of prewhitening various types of data. Table 1 shows our results on 162 samples spectrally analyzed. It shows average slopes (from the fitted lines), standard deviations (s.d.) of these slopes, and s.d. of the log amplitudes of the spectra at the "center of gravity" wavenumber $1.77 \times 10^{-2}$ rad/m. The results for data with “input slopes” $-2.5$, $-3.0$, and $-3.5$ are shown for cases with and without prewhitening (PW). These results correspond to Figures 3, 4, 5, 9 (top), 10 (top), and 11 (top). We can see from these observations that (1) as input slope steepens, the s.d. of the spectral log amplitudes become larger when there is no PW, whereas (2) with PW, all s.d. are reduced to the same lower level. Finally, (3) when PW is used, the resulting output average slope agrees with the input simulated slope, whereas without PW, as was already mentioned, there is an increasing error in the average slope as input slope increases in steepness. These observations are consistent with the fact that the SD artifact is caused by leakage. As was mentioned above, this is because
first differencing causes data of spectral slope $-2$ to become data of zero slope. Data associated with a $-3$ slope become $-1$ slope (and so on), and the latter seems to be sufficiently white, in the case of 128 points, to do little damage.

It is also important to recall that without PW, one obtains an incorrect average slope; thus averaging over many spectra alone does not combat leakage artifacts. Prewhitening is required.

It is interesting to note that in the case where no PW is used, the SD of the slopes increased with “input spectral steepness.” In contrast, when PW is used, the s.d. of the slopes observed are independent of “input spectral steepness” and are reduced in size.

It is, however, both surprising and very significant that when a single slope of $-3$ is input to the model, the analyzed data (even with PW) have slopes that range from $-1$ to almost $-5$! This means that even when the 128 point data samples are analyzed properly, they have slopes ranging over $-3 \pm 2$ for a 162 point sample, and thus such samples give an extremely crude estimate of the “actual” spectral slope involved. The error bars on the experimental spectral slopes do not reflect this reality. It is interesting to note, in contrast, that for the 1000 point spectrum of Figure 14 (top) (see also Table 2) the spread in slopes has decreased to $\pm 0.5$ instead of $\pm 2$.

Table 2 shows a comparison between results from data with 128 and 1000 points. In all cases, a $-3$ slope input spectrum was used. In contrast to all previous cases, no “lag window” is employed here. The comparison between the unprewhitened 128 point case and the unprewhitened 1000 point case shows that by going from 128 to 1000 points, one reduces the log amplitude s.d. by about as great a factor as using PW on the 128 point case. However, even when one has 1000 points, we see that PW causes a further significant reduction in this s.d.

Finally, we consider Table 3, which shows 128 point data sets with input slopes of $-3$. This table compares results between cases where (1) there is no PW, (2) there is PW, (3) there is Hann data tapering (an alternative leakage reduction technique), and finally, (4) there is both PW and data tapering at the same time. (The reader is reminded that data tapering with windows has a different function from “lag window” tapering. The latter “windows” the autocorrelation in order to reduce spectral variability. For explanation, see Percival and Walden [1993].)

As can be seen from Table 3, the behavior of the log amplitude s.d. indicates that the best procedure to reduce leakage and the associated “variability artifact” is prewhitening. The Hann data taper (the best window to reduce leakage) does indeed have a helpful effect, but it is about 15% less helpful than PW in the present example. Finally, it should be noted that adding a Hann data taper to PW does not improve the situation at all! (The small number of data points combined with the number of lost degrees of freedom from use of the window probably explains this unexpected outcome.) Another interesting observation is that the slope s.d. for no PW is reduced for PW, whereas data tapering increased it.

Incidentally, it should be mentioned that we have been con-
Figure 9. (top) Scatterplot for input spectral slope of (−3.5) but with prewhitening. (bottom) Chi squared distributions. Note that with prewhitening the theoretical and experimental curves agree.
Figure 10. (top) Scatterplot for input spectral slope of \((-3.0)\) with prewhitening. (bottom) Chi squared distributions.
Figure 11. (top) Scatterplot for input spectral slope of \((-2.50)\) with prewhitening. (bottom) Chi squared distributions.
Figure 12. (top) Scatterplot for input spectral slope of (-3.3) without prewhitening. Fitted line agrees with balloon observations. (bottom) Chi squared distributions.
Figure 13. (top) Scatterplot for input spectral slope of $(-3.3)$ but with 1000 points of simulated data instead of 128 points above. No prewhitening used. (bottom) Chi squared distributions.
Figure 14. (top) Scatterplot for input spectral slope of (−3.3) and with 1000 points of data but with prewhitening. Note that now the fitted line here is essentially horizontal. (bottom) Chi squared distributions.
considering the standard deviations of the spectral log amplitudes. The standard deviations of the amplitudes would therefore depend on the amplitudes as a consequence; that is, logarithmic variations represent fractional variations.

### 6. Conclusions

In this paper we considered power spectral densities, i.e., PSDs, a new artifact, and the possible impact of the latter on the published experimental conclusions from PSDs in relation to atmospheric gravity wave theories. To facilitate this, we have presented a new and practical method to simulate data of given PSD properties. This method does not suffer from the flaws that are present in some previously published methods [e.g., Owens, 1978; Fairall and White, 1991]. The new, as yet unrecognized, artifact that we found consisted of a dramatic increase in spectral log-amplitude variability when “red” stochastic data are analyzed. We found that when leakage is sufficiently reduced by some method of prewhitening, the artifact is eliminated. Data tapering with a Hann window did not work as effectively as prewhitening in our case.

More specifically, we have shown that in the case of gravity wave data, when steep negative slopes occur (in the range $-2.5$ to $-3.5$), the procedure of section 3 above will give valid results. The prewhitening procedure of (21), followed later by the postcoloring procedure of (25), was found to eliminate the leakage artifact, which is the subject of this paper. An alternative procedure, which we found to be very successful for the type of data under consideration, was published by Dewan et al. [1988]. The reader is now in a good position to judge which of any methods of PSD analysis would be best to use for his or her type of data set by means of simulations.

In the general case of PSD analysis of arbitrary data, which is beyond the scope of this report, we refer the reader to Percival and Walden [1993], which we agree is the best book on how to carry this out, for example. The latter type of PSD (also called MEM or maximum entropy method), by virtue of its “leakproof” nature, will show immediately by this comparison whether or not the FFT PSD is showing leakage effects. (AR spectra have different artifacts of their own.) Another hint to detect leakage is described on p. 288 of Percival and Walden [1993]. We would like to conclude this paper by saying that we have

---

### Table 1. Results on 162 Samples Spectrally Analyzed, PW versus not PW

<table>
<thead>
<tr>
<th>Theory log$_{10}$ PSD Input</th>
<th>Neg. of Slope</th>
<th>Number of Points</th>
<th>Data Taper?</th>
<th>Prewhiten?</th>
<th>Avg log$_{10}$ PSD Amp.</th>
<th>Avg PSD Neg. Slope</th>
<th>s.d. log$_{10}$</th>
<th>s.d. Slope</th>
<th>Figure</th>
<th>Hann Log Window?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>2.5</td>
<td>128</td>
<td>not PW</td>
<td></td>
<td>1.01</td>
<td>2.40</td>
<td>0.139</td>
<td>0.468</td>
<td>5</td>
<td>Y</td>
</tr>
<tr>
<td>1.88</td>
<td>3.0</td>
<td>128</td>
<td>not PW</td>
<td></td>
<td>1.99</td>
<td>2.72</td>
<td>0.215</td>
<td>0.515</td>
<td>3</td>
<td>Y</td>
</tr>
<tr>
<td>2.75</td>
<td>3.5</td>
<td>128</td>
<td>not PW</td>
<td></td>
<td>3.03</td>
<td>2.88</td>
<td>0.325</td>
<td>0.624</td>
<td>4</td>
<td>Y</td>
</tr>
<tr>
<td>1.00</td>
<td>2.5</td>
<td>128</td>
<td>PW</td>
<td></td>
<td>0.922</td>
<td>2.45</td>
<td>0.0943</td>
<td>0.461</td>
<td>11 (top)</td>
<td>Y</td>
</tr>
<tr>
<td>1.88</td>
<td>3.0</td>
<td>128</td>
<td>PW</td>
<td></td>
<td>1.80</td>
<td>2.95</td>
<td>0.0944</td>
<td>0.459</td>
<td>10 (top)</td>
<td>Y</td>
</tr>
</tbody>
</table>

PSD, power spectral density; PW, prewhitening.

---

### Table 2. Comparison Between Results from Data With 128 and 1000 Points

<table>
<thead>
<tr>
<th>Theory log$_{10}$ PSD Input</th>
<th>Neg. of Slope</th>
<th>Number of Points</th>
<th>Data Taper?</th>
<th>Prewhiten?</th>
<th>Avg log$_{10}$ PSD Amp.</th>
<th>Avg PSD Neg. Slope</th>
<th>s.d. log$_{10}$</th>
<th>s.d. Slope</th>
<th>Figure</th>
<th>Hann Log Window?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.40</td>
<td>3.3</td>
<td>128</td>
<td>not PW</td>
<td></td>
<td>2.60</td>
<td>2.84</td>
<td>0.2780</td>
<td>0.574</td>
<td>12 (top)</td>
<td>Y</td>
</tr>
<tr>
<td>2.40</td>
<td>3.3</td>
<td>1000</td>
<td>not PW</td>
<td></td>
<td>2.39</td>
<td>3.12</td>
<td>0.0803</td>
<td>0.254</td>
<td>13 (top)</td>
<td>Y</td>
</tr>
<tr>
<td>2.40</td>
<td>3.3</td>
<td>1000</td>
<td>PW</td>
<td></td>
<td>2.32</td>
<td>3.29</td>
<td>0.0348</td>
<td>0.183</td>
<td>14 (top)</td>
<td>Y</td>
</tr>
</tbody>
</table>
never, in our experience, found it possible to do valid PSD analysis without doing appropriate tests with simulated data. (The idea is to see if the output of the analysis matches the known input based upon what a priori knowledge we have of the data. The latter can be based on low-order AR analysis as described by Percival and Walden [1993]). In the present context this raises troubling questions about some of the PSDs published in the literature that may have been done without testing but instead with full reliance on numerical recipes and perhaps little knowledge of the many possible artifacts (again, see Percival and Walden [1993]).

Acknowledgments. We would like to thank C. Gardener, B. Sears, and J. Brown and, most especially, G. Nastrom and T. Van Zandt for helpful discussions in regard to this work. This work was sponsored under two AFSOR projects, namely the SOAR and ABL projects.

References


E. M. Dewan, Air Force Research Laboratory, Space Vehicles Directorate, Hanscom AFB, MA 01731-3010. (edewan@plh.af.mil)
N. Grossbard, Boston College, Institute for Scientific Research, Chestnut Hill, MA 02467-3862.

(Received March 29, 1999; revised October 5, 1999; accepted October 14, 1999.)

Table 3. Data Sets of 128 points With Input Slopes of $-3$, Comparing Effects of PW and Data Tapers

<table>
<thead>
<tr>
<th>Theory</th>
<th>Neg. of</th>
<th>Number</th>
<th>Data</th>
<th>Avg log$_{10}$</th>
<th>Avg PSD</th>
<th>s.d. log$_{10}$</th>
<th>s.d.</th>
<th>Hann</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSD Input</td>
<td>Slope Input</td>
<td>Points</td>
<td>Taper? Prewhiten?</td>
<td>PSD Amp.</td>
<td>PSD Amp.</td>
<td>Slope</td>
<td>Slope</td>
<td>Window?</td>
</tr>
<tr>
<td>1.88</td>
<td>−3.0</td>
<td>128</td>
<td>PW</td>
<td>1.66</td>
<td>−3.0</td>
<td>0.109</td>
<td>0.569</td>
<td>···</td>
</tr>
<tr>
<td>1.88</td>
<td>−3.0</td>
<td>128</td>
<td>not PW</td>
<td>1.89</td>
<td>−2.63</td>
<td>0.240</td>
<td>0.604</td>
<td>···</td>
</tr>
<tr>
<td>1.88</td>
<td>−3.0</td>
<td>128</td>
<td>Hann data taper</td>
<td>1.65</td>
<td>−3.09</td>
<td>0.132</td>
<td>0.655</td>
<td>···</td>
</tr>
<tr>
<td>1.88</td>
<td>−3.0</td>
<td>128</td>
<td>PW + Hann data taper</td>
<td>1.64</td>
<td>−3.07</td>
<td>0.133</td>
<td>0.655</td>
<td>···</td>
</tr>
</tbody>
</table>