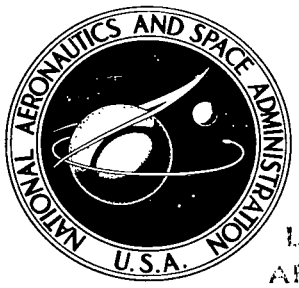
NASA TECHNICAL NOTE



NASA TN D-8288

AN ASSESSMENT OF PREWHITENING IN ESTIMATING POWER SPECTRA OF ATMOSPHERIC TURBULENCE AT LONG WAVELENGTHS

Samuel R. Keisler and Richard H. Rhyne

Langley Research Center
Hampton, Va. 23665

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • NOVEMBER 1976

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA TN D-8288		

4. Title and Subtitle	5. Report Date
AN ASSESSMENT OF PREWHITENING IN ESTIMATING POWER SPECTRA OF ATMOSPHERIC TURBULENCE AT LONG WAVELENGTHS	November 1976
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Samuel R. Keisler and Richard H. Rhyne	L-10502
	10. Work Unit No.
9. Performing Organization Name and Address	743-01-12-02
NASA Langley Research Center Hampton, VA 23665	11. Contract or Grant No.
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Technical Note
National Aeronautics and Space Administration Washington, DC 20546	14. Sponsoring Agency Code

15. Supplementary Notes

Samuel R. Keisler: SDC Integrated Services, Inc.
Richard H. Rhyne: Langley Research Center.

16. Abstract

Synthetic time histories were generated and used to assess the effects of prewhitening on the long wavelength portion of power spectra of atmospheric turbulence. Prewhitening is not recommended when using the narrow "spectral windows" required for determining power spectral estimates below the "knee" frequency, that is, at very long wavelengths.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Atmospheric turbulence Power spectra Prewhitening Turbulence models Stochastic processes	Unclassified - Unlimited Subject Category 46

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	54	$4.25

* For sale by the National Technical Information Service, Springfield, Virginia 22161

AN ASSESSMENT OF PREWHITENING IN ESTIMATING POWER SPECTRA OF ATMOSPHERIC TURBULENCE AT LONG WAVELENGTHS

Samuel R. Keisler* and Richard H. Rhyne
Langley Research Center

SUMMARY

An assessment is given of the effects of prewhitening on determination of power spectra of atmospheric turbulence at long wavelengths. In a computer experiment, a synthetic time history was generated by combining sine waves of random phase and frequency (but uniformly distributed over the frequency range of interest) with amplitudes adjusted to produce a power spectrum of known shape approximating that of atmospheric turbulence. The synthetic time history was then used to assess bias errors in power spectra computed by three different algorithms implemented by the fast Fourier transform.

Prewhitening is unnecessary when using the narrow "spectral windows" required for determining power spectra of atmospheric turbulence below the "knee" frequency, or at very long wavelengths. Several wide spectral window cases are also included to assess the effect of first-difference prewhitening on data where the first spectral estimate is above the knee frequency.

INTRODUCTION

The generally accepted mathematical model of the power spectral density function for atmospheric turbulence (the Von Kármán

*SDC Integrated Services, Inc.

model; see ref. 1) describes the amplitude of the spectrum as nearly constant at very low frequencies, varying inversely with the five-thirds power of frequency at the higher frequencies. The transition between these frequency regions is often referred to as the "knee" and occurs at a frequency that is dependent on the integral scale of turbulence (L). The knee frequency, and therefore the scale of turbulence, is significant with respect to the calculated motion and load responses of aircraft (ref. 1).

The adequacy of the Von Kármán spectrum as a model of atmospheric turbulence has not been assessed sufficiently. Although the shape of the spectrum at high frequencies has been shown to agree with measured turbulence spectra, there is still uncertainty about the representation by the Von Kármán spectrum at frequencies associated with and below the knee of the theoretical spectrum. This is caused, to a large extent, by difficulties in accurately determining values in the low frequency region of spectra of atmospheric turbulence from flight measurements. These difficulties include (1) obtaining adequate length samples, (2) inaccuracies in measurements of airplane motions that are required to extract the turbulence velocities, and (3) errors introduced by the conversion of the measured time histories of the velocities to power spectral density functions. Efforts are underway to reduce errors from the first two sources. This report is concerned with the third problem cited.

A significant source of error in the determination of experimental spectra is the bias error associated with data processing. The bias error is negligible for spectra of essentially constant amplitude, but it can produce an unacceptably large distortion in spectra exhibiting large changes in amplitude. Therefore, the high-frequency region of the turbulence spectrum, where the amplitude varies inversely with the five-thirds power of frequency, is susceptible to bias error. Press and Tukey (ref. 2) introduced a conditioning process called prewhitening to reduce the distortion of turbulence spectra substantially (that distortion caused by

data processing) at the higher frequencies since the higher frequency range was the region of particular interest at the time. The prewhitening procedure, which has become routine, renders the high-frequency range of the spectrum to nearly constant amplitude during processing. The effect is then removed (postdarkened) for final data presentation. If, however, the actual turbulence spectrum in the low-frequency region is nearly constant, as suggested by the Von Kármán spectrum model, the prewhitening process results in a bias distortion of the spectrum in this low-frequency range. Current interest in improved definition of measured spectra at low frequencies prompted a reassessment of the use of the prewhitening process and other aspects of the bias error problem.

The present computer study[1] is similar in approach to the study described in reference 3 in that spectra are estimated by the digital processing of a synthetic random time function that simulates the velocities of atmospheric turbulence. In contrast to reference 3, which examined the effects of various methods of prewhitening, the present study compares spectra obtained with and without prewhitening, with emphasis on the effects at low frequencies. The method used in this study to generate the synthetic random time function differs from the method used in reference 3; the reasons for this different method are discussed in a later section. The data processing methods consist of three commonly used fast Fourier transform (FFT) algorithms: (1) Blackman-Tukey, (2) ensemble averaging, and (3) frequency averaging.

SYMBOLS

Values are given in both SI and U.S. Customary Units. The measurements were made in U.S. Customary Units.

---

[1]Recommended by Mr. Allan Piersol (Bolt, Beranek, and Newman, Inc., Canoga Park, CA).

B_e equivalent spectral bandwidth, Hz

$B_{e_k'}$ dimensionless equivalent spectral bandwidth, $\frac{2\pi L}{V} B_e$

f frequency, Hz

G_i smoothed power spectral density estimate at frequency $i \Delta f$

G'_i raw power spectral density estimate at frequency $i \Delta f$

$H(f)$ transfer function

i, j, k, n index, integer variable

K number of points, including zeros, transformed by fast Fourier transform (power of 2)

k' dimensionless wave number, $\frac{2\pi fL}{V}$

L integral scale of turbulence

ℓ number of nonnegative lags of time-sampled autocorrelation function

M number of contiguous raw power spectral density estimates which are used to compute each point of frequency averaged power spectral density; or number of signal segments used in computing ensemble average power spectral density

N	integer constant, i.e., final integer in function sequence
P	time length of function (or segment of function) being transformed
s_i	frequency-sampled spectral estimate at frequency $i\ \Delta f$
T	turbulence sample length, sec
t	time, sec
V	velocity
x	random time function
x_i	time function sampled at time $i\ \Delta t$
y	prewhitened random time function
y_i	prewhitened time function sampled at time $i\ \Delta t$
Δf	frequency interval, Hz
Δt	time sampling interval, sec
θ	phase angle, rad
τ	lag time
τ_m	maximum lag time
Φ	power spectrum as function of f

$\tilde{\Phi}$ power spectrum as function of k'

Φ_T true Dryden spectrum undistorted by bias error

Φ_x power spectrum of postdarkened time function

Φ_y power spectrum of prewhitened time function

ω angular or circular frequency, $2\pi f$

ω_1 circular frequency integration variable

APPROACH

Since the primary purpose of this study is to assess the value of prewhitening in reducing bias error in estimated spectra at low frequencies, a review of bias error is appropriate. This review includes: (1) properties related to the general digital processing of a random time function, (2) effects at high and low frequencies, and (3) a brief description and discussion of the prewhitening process. The review is followed by descriptions of the three digital algorithms for estimating power spectral densities and their relative bandwidth properties. The properties of a synthetic random time function needed to represent atmospheric turbulence velocities are then discussed, and the two functions used in numerical calculations are described. The conditions selected for estimating power spectral densities are stated and results are discussed.

BIAS ERROR

Each of the power spectral density algorithms considered here uses the Fourier transform of a time function of finite length. The following discussion is equally applicable to the Fourier transform of a finite length autocorrelation function as well as a finite

length time function. Further details are given on page IVC:19 of reference 2. The finite time length of a time function affects the measured spectral function and can be considered the product of an infinite length function (given the subscript "true") and a "boxcar" function (fig. 1(a)). The measured and the true time functions are equal when $\frac{-P}{2} < t < \frac{P}{2}$. The measured time function is equal to zero for all other values of time. The Fourier transforms of the two time functions are related by the convolution integral

$$\Phi_{meas}(\omega) = \int_{-\infty}^{\infty} \Phi_{true}(\omega_1) P \frac{\sin\left[(\omega - \omega_1)\frac{P}{2}\right]}{(\omega - \omega_1)\frac{P}{2}} d\omega_1 \qquad (1)$$

where $\Phi_{true}(\omega)$ is the "true" spectrum for limit $P \to \infty$ and the $\frac{\sin x}{x}$ term is the spectral "window." (See fig. 1(b).) The difference between $\Phi_{meas}(\omega)$ and $\Phi_{true}(\omega)$ is defined as the bias error. (This example is applied to the boxcar window; however, the same relation would hold for other windows.) Convolution can be understood heuristically by considering its effect on an arbitrary spectral estimate $\Phi_{meas}(\omega)$, according to the following procedure: Shift the spectral window so that it is centered at ω; form the product of the true spectrum Φ_{true} and the shifted window; then $\Phi_{meas}(\omega)$ is the integral of this product.

The strongest objection to the boxcar window concerns "leakage through the side lobes" (fig. 1(b)); the side lobes are those portions of the window between $\frac{1}{P}$ and $\frac{2}{P}$, between $\frac{2}{P}$ and $\frac{3}{P}$, and so forth. (The portion between $\frac{-1}{P}$ and $\frac{1}{P}$ is called the "main lobe.") When the process of convolution is kept in mind, it can be seen that any portion of the true spectrum which is multiplied by a side-lobe peak during convolution can have a significant effect upon the estimate.

This deficiency in the boxcar window is largely overcome in windows, such as the Hann window (ref. 4), shown in figures 2(a) and 2(b), which taper the ends of the function. The Hann side lobes are greatly reduced in relation to the boxcar side lobes.

Figure 3 illustrates a typical bias error encountered in estimating the power spectrum. The true spectrum together with the shifted spectral window are shown with linear scale on each axis. The true spectrum $\Phi_{true}(\omega)$ shown in figure 3 is an approximation of the Von Kármán turbulence model, and is discussed later in more detail. It can be visualized that the magnitude and sign of the bias error depend primarily upon the broadness and general shape of the main lobe of the spectral windown and its position along the frequency axis with respect to the true spectrum. In general, the peak of the true spectrum is underestimated, and all other values overestimated.

Each of the three power spectral density algorithms is associated with a different variant of the fundamental spectral window described above. The differences among these variants are affected very little by the time window chosen, e.g., Hann, Hamming, or Parzen (see ref. 4). For this reason, only the Hann time window is considered here.

PREWHITENING

The prewhitening process consists of filtering the random time function so as to produce a nearly flat spectrum at frequencies above the knee. (The purpose, as stated earlier, is to reduce or eliminate the bias error caused by the window convolution illustrated in fig. 3.) The spectral estimates are then "postdarkened" to correct for the initial prewhitening. The most commonly used prewhitening process is the first-difference method

$$y(t) = x(t) - x(t - \Delta t)$$

or in digitized sample form

$$y_i = x_i - x_{i-1} \tag{2}$$

The spectrum estimated from the prewhitened function by a particular digital algorithm is then postdarkened as follows before final data presentation

$$\Phi_x(f) = \frac{1}{|H(f)|^2}\,\Phi_y(f) \tag{3}$$

where

$$\frac{1}{|H(f)|^2} = \frac{1}{2(1 - \cos 2\pi f\ \Delta t)}$$

(See ref. 2, or appendix E of ref. 8.)

The true spectrum $\Phi_T(f)$ of figure 3 is shown in figure 4 as $\Phi_y(f)$ after the first-difference prewhitening filter has been applied, i.e., $\Phi_y(f) = \Phi_T(f)\,|H(f)|^2$. The knee frequency of figure 3 is indicated by an arrow on the abscissa of figure 4.

(The frequency scale in fig. 4(b) is exaggerated in comparison to the scale of fig. 4(a) to show more clearly the low-frequency region.) It is apparent that the bias error will be positive at the knee frequency and below, and will become larger as the spectral window becomes broader. At frequencies well above the knee, as the prewhitened spectrum becomes more nearly flat, the bias error will become negligible, even though the spectral window is quite broad.

POWER SPECTRAL DENSITY ALGORITHMS

The turbulence time histories are sampled in time at interval Δt. The time function is denoted by x_i and is the value of the function sampled at time $i\ \Delta t$ (where $i = 1, 2, \ldots, N$); then discrete spectral estimates s_i corresponding to harmonic frequencies $i\ \Delta f$ are computed by the FFT $\left(\text{where } i = 0, 1, 2, \ldots, \frac{N}{2}\right.$ and $\left.\Delta f = \frac{1}{N\ \Delta t}\right)$. The time length of the function (or segment of function) being transformed is $P = N\ \Delta t$; so

$$\Delta f = \frac{1}{P} = \frac{1}{N\ \Delta t} \tag{4}$$

where Δf is the frequency spacing of the original "raw" estimates. Each of the power spectral density algorithms is now examined together with its associated spectral window.

Blackman-Tukey

The technique for computing the autocorrelation function with Fourier transforms rather than with lag products as used

in the past is given in reference 4 (pp. 165 and 166). The Blackman-Tukey FFT power spectral density algorithm is outlined briefly in the following eight steps.

(1) Let K, which must be a power of two, be the number of points being transformed. Let ℓ be the number of nonnegative lags required for the autocorrelation function. Only $K - \ell$ data points can be used, since the last ℓ values of the points being transformed, at least, must be zeroed to avoid the circular autocorrelation effect (ref. 4, pp. 123 and 124). The mean of the data function is removed and the correct number of zeros is added.

(2) Optional prewhitening: the first difference filter is applied (eq. (2)).

(3) The FFT is computed to yield the complex amplitude spectrum.

(4) The "raw" power spectrum is computed as the scaled square modulus of the result of step (3).

(5) The inverse FFT is computed to produce the autocorrelation function.

(6) The Hann window is applied.

(7) The FFT is computed to obtain the power spectrum, which is then scaled to obtain the power spectral density.

(8) Postdarkening (if option (2) is selected) is applied as indicated by equation (3).

The autocorrelation function computed in step (5) is defined for $-\tau_m \leq \tau \leq \tau_m$ where τ is lag time and $\tau_m = (\ell - 1)\Delta t$. The quantity ℓ is adjusted to truncate the autocorrelation function at an appropriate maximum lag τ_m which results in the desired spectral window width B_e. For convenience in the FFT procedure, $\ell - 1$ should also be a power of two. If τ_m is set equal to $\frac{P}{2}$ in equation (4), $\Delta f = \frac{1}{2\tau_m}$. When the Hann window shown in figure 2 is used in step (6), then the spectral windows for two adjacent

spectral estimates at frequencies $i\,\Delta f$ and $(i + 1)\Delta f$ will be as shown in figure 5. The mechanics of the Blackman-Tukey algorithm are such that the frequency spacing of the final power estimates Δf is always $\frac{1}{2\tau_m}$, and the effective spectral window width B_e is $\frac{1}{\tau_m}$ or $2\,\Delta f$. A rather large degree of window overlap for adjacent spectral estimates is shown in figure 5. It is apparent, therefore, that such estimates are not strictly independent. For this reason, it is sometimes advocated that every other estimate be discarded when employing this algorithm.

Consider now the estimate s_0 at zero frequency. The zero frequency component of the signal (i.e., mean) was removed in step (1). Therefore it appears that s_0 is an estimate of the power at positive frequencies. For this reason, s_0 is (somewhat arbitrarily) displayed at frequency $\frac{\Delta f}{4}$ in this paper.

Ensemble Averaging

This algorithm is described in detail in reference 5. A brief outline follows.

(1) The time function is partitioned into M segments of equal length, with the number of points in each segment a power of two. Steps (2) to (6) must be performed independently for each segment.

(2) The mean of the function is removed.

(3) (Optional prewhitening) The first difference filter is applied.

(4) The side-lobe suppression window is applied.

(5) The FFT is computed to obtain the complex amplitude spectrum.

(6) The power spectral density estimates are computed as the scaled square modulus of the amplitude spectrum.

(7) The M segment power spectral density functions are averaged.

(8) (Optional) Postdarken.

In reference 5, the observation is made that the expected value of the spectrum at each point is the convolution of the true spectrum with the square of the spectral window if the time window is even (a proof of this is given in ref. 6). All commonly used windows (including the Hann window) are even.

The squared Hann window is shown in figure 6. It is the effective window in the limit as the length of the signal, and therefore the number of segments used, increases without bound. For this reason, the squared spectral window will be referred to here as the "expected window."

Figure 6 shows that the ensemble average window bandwidth B_e, that is, the width of the main lobe at the half-power point, is approximately 1.5 Δf. Let $\frac{K}{M}$ be the number of points per segment; then when $\frac{K}{M} = N$ (from eq. (4)), the spacing of the power estimates is Δf, equal to $\frac{M}{K\ \Delta t}$ or

$$B_e \approx 1.5\ \Delta f \approx \frac{1.5M}{K\ \Delta t}$$

Note that in this figure contiguous windows overlap less than for the Blackman-Tukey algorithm; therefore, contiguous ensemble average estimates are statistically more independent of each other.

Frequency Averaging

This algorithm is a very simple modification of the ensemble averaging algorithm.

(1) Perform steps (2) to (6) of the ensemble averaging method for the entire time function consisting of K points. Denote the resulting spectrum G'_i, where $i = 0, 1, \ldots, \frac{K}{2}$.

(2) The smoothed power spectral density estimates, here denoted G_j, are computed as the means of M contiguous values of G'_i as follows:

$$G_j = \frac{G'_{k+1} + G'_{k+2} + \ldots + G'_{k+M}}{M}$$

and

$$j = 1, 2, \ldots, \frac{K}{2M}; \quad k = (j - 1)M$$

(3) (Optional) Postdarken as above.

The expected window is now the sum of M contiguous windows, each as shown in figure 6 (see ref. 7, p. 296). The frequency average expected window for $M = 16$ is shown in figure 7; the estimate is displayed at the midpoint of the window, i.e., the jth point of the spectrum is displayed at frequency $\frac{(2j - 1)M + 1}{2K\ \Delta t}$. The final frequency spacing of the power estimates Δf is then $\frac{M}{K\ \Delta t}$, and the bandwidth B_e is approximately $\frac{M + 0.5}{K\ \Delta t}$. As can be

seen in figure 7, the expected window is now essentially rectangular, with virtually no overlap of contiguous windows.

In comparison to ensemble averaging, frequency averaging has this additional advantage: for the same frequency spacing of the final power estimates (and approximately the same bandwidth B_e), the lowest frequency analyzed is lower by approximately a factor of two. The lowest frequency estimate appears at $\frac{M + 1}{2K \Delta t}$ for frequency averaging, as opposed to $\frac{M}{K \Delta t}$ for ensemble averaging.

SYNTHETIC RANDOM TIME FUNCTIONS

For this specific study, the synthetic random time function must possess the significant characteristics of atmospheric turbulence velocities. These include randomness, a nearly Gaussian probability density function, and a power spectral density representative of that for turbulence. The power spectral density was chosen to be the Dryden spectrum (see ref. 8) to parallel the study of reference 3. The differences between the Dryden and Von Kármán spectra are not considered to be significant for the present purposes. The Dryden spectrum is described by

$$\tilde{\Phi}(k') = \frac{1}{\pi} \frac{1 + 3k'^2}{(1 + k'^2)^2} \tag{5}$$

where

$$k' = \frac{2\pi f L}{V}$$

(sometimes referred to as "dimensionless wave number"), and where

L integral scale of turbulence

V velocity of aircraft

f frequency, Hz

In terms of temporal frequency

$$\Phi(L,V,f) = \frac{2L}{V} \frac{1 + 3\left(\frac{2\pi fL}{V}\right)^2}{\left[1 + \left(\frac{2\pi fL}{V}\right)^2\right]^2} \qquad (6)$$

The synthetic function, of practical necessity, is synthesized from a finite set of numbers and therefore, cannot completely describe a power spectral density that is continuous in frequency. Consequently, it is necessary to analyze any particular synthetic signal for adequacy in this respect. This was done for the signal used in reference 3 (designated herein as Signal B) and details are given in the appendix. Signal B is a model of a process whose frequency components are harmonics of the fundamental frequency $\frac{1}{P}$. It is therefore not a good representation of atmospheric turbulence, whose frequency components are distributed continuously over all frequencies. The frequency components of Signal B coincide precisely with the nulls of the boxcar window and with the center of the main lobe. Consequently, Signal B can produce no bias error with the boxcar window. This effect is considered unrealistic with respect to simulating atmospheric turbulence.

To model the continuous frequency distribution of turbulence more accurately, an alternate synthetic time function, designated Signal A, was derived for use in this study. Details concerning this signal are also given in the appendix.

$$\text{Signal A} = \sum H(f) \sin (2\pi f_R n \Delta t + \theta_R)$$

where both f_R and θ_R are randomly picked values in each of 1250 frequency intervals. Thus, spectral values do not necessarily coincide with boxcar window nulls, and the approximation of atmospheric turbulence is improved. A short segment of the time function is shown in figure 8; its appearance is very much like that of actual turbulence time-history measurements. The cumulative distribution function, shown in figure 9, is very nearly Gaussian, as desired, and as expected from the procedure used.

CONDITIONS CONSIDERED FOR SPECTRAL ESTIMATION

To determine a practical range of bandwidth B_e which should be investigated for bias error effects, constraints imposed by sample-length limitations of actual turbulence data must be considered. It would, of course, be desirable to have the spectral estimates very closely spaced (small B_e) in the frequency region approaching zero (see fig. 3) to define a knee, if one exists. As the integral scale of turbulence L becomes larger in the mathematical model, the knee frequency moves closer to zero frequency, and in fact might be very close to zero frequency for certain meteorological conditions. For example, if L is 762 m (2500 ft) in the Von Kármán turbulence model (a present turbulence design requirement; see ref. 9) and the velocity of the airplane is 183 m/s (600 ft/sec), the knee peak appears at 0.017 Hz. Larger L values result in lower knee frequencies.

As the bandwidth is made smaller (for a given sample length of actual turbulence data), the random fluctuations of the power

estimates become larger, eventually making it impossible to determine whether a knee is present. Past experience and a calculation of the magnitude of the random power fluctuations based upon so-called "statistical degrees of freedom" (see ref. 7) indicate that approximately 30 degrees of freedom is as low as can be tolerated.

According to reference 7 (p. 219),

$$\text{Degrees of freedom (d.o.f.)} = 2B_eT \tag{7}$$

where B_e is bandwidth in Hz, and T is total sample length in seconds. Past experience with turbulence data collection indicates that usable sample lengths of more than 10 minutes duration are not readily obtainable, particularly at the higher altitudes. If we assume the maximum practical sample length to be 10 minutes (or 600 sec), and that 30 d.o.f. can be tolerated, then B_e is 0.025 Hz.

Based on the preceding discussion, indications are that although narrower bandwidths for better spectral resolution at very low frequencies would be desirable, bandwidths smaller than about 0.025 Hz are not practical because of sample length limitations. The present bias error investigation will therefore be confined to bandwidths on the order of 0.025 Hz and greater. (It should be noted that random error is only of indirect concern in this study, and that it has, in fact, been minimized by the synthetic time functions so as not to obscure the bias error under study.)

Since this study was initiated by following the procedure of reference 3, the numbers actually used in equation (6) to generate the synthetic signals were L = 191 m (627 ft) and V = 183 m/s (600 ft/sec) to conform with reference 3. The synthetic time function values were spaced at an interval of Δt = 0.04 sec, and this spacing provided a Nyquist frequency of the resulting power spectral density of 12.5 Hz, also in accordance with reference 3.

A more general, and possibly more useful, procedure would have been to generate the spectral results in terms of dimensionless wave number k' (see eq. (5)) rather than in terms of temporal frequency. The correct shape of the spectrum is all that is required, however, and the exact location of the knee peak along the frequency axis will not affect the magnitude of the bias error obtained. If the location of the knee peak is known in advance, and/or if an estimate of bias error for a specific L, B_e, or V at a particular frequency relative to the knee peak is needed, then k' units can be helpful. A k' scale has therefore been included on all the spectral plots presented. Also B_e is given in k' units as well as in Hz, and is then designated $B_{e_{k'}}$.

The synthetic time function was processed as if it were measured data. Spectra were estimated by the three algorithms with and without first-difference prewhitening. As indicated, the minimum bandwidth desired was approximately 0.025 Hz. The nearest bandwidths readily obtainable with the three algorithms were B_e = 0.0244, 0.0366, and 0.0260 Hz for Blackman-Tukey, ensemble averaging, and frequency averaging (or $B_{e_{k'}}$ = 0.160, 0.240, and 0.171), respectively. These values were determined as discussed under power spectral density algorithms, by use of 1024 nonnegative lags for Blackman-Tukey, 15 time segments each containing 1024 data points for ensemble averaging, and frequency averaging over 16 contiguous raw estimates obtained by transforming 15 871 data points plus 513 zeros to make an even power of two. The Δf frequency spacing was thus 0.0122 Hz for Blackman-Tukey and 0.0244 Hz for ensemble averaging and frequency averaging.

Additional comparisons between prewhitened and non-prewhitened spectra were made with the Blackman-Tukey algorithm for an intermediate B_e of 0.195 Hz ($B_{e_{k'}}$ = 1.28) employing 128 lags, and for a very wideband B_e of 0.391 Hz or $B_{e_{k'}}$ of 2.56 (64 positive lags). The wideband spectra simulated results obtained in earlier years before FFT was available.

RESULTS AND DISCUSSION

In the following figures, the power spectral density is plotted along the vertical axis, and the frequency in Hz and in k' units is plotted along the horizontal axis, both on logarithmic scales. The true power spectral density $\Phi_T(f)$ is modified by a multiplication factor selected so that the area under $\Phi_T(f)$ will be equal to the variance of the synthetic random time function. The obtained $\Phi_T(f)$ is superimposed on the plot of each computed power spectral density. In each case, the "true" spectrum is labeled "reference" and the spectrum calculated from the time history is labeled "computed."

The results for the minimum bandwidth case, presented in figures 10, 11, and 12, indicate that all three algorithms define the knee adequately. There is an indication of bias error (a slight overestimation in the estimates nearest zero frequency) only in the spectrum obtained with the Blackman-Tukey algorithm (fig. 10).

With all other parameters held fixed, first-difference prewhitening and postdarkening were applied. In each case the effect (see figs. 13, 14, and 15) was an overestimation at the lowest frequencies and no significant improvement elsewhere. The curiously large difference between the amount of distortion obtained by using the Blackman-Tukey algorithm and the amount of distortion obtained by using the direct-transform algorithms needs an explanation. To rule out the possibility of error in the Blackman-Tukey portion of the computer program, the synthetic signal was processed through a lag-product Blackman-Tukey program (see ref. 10), and the results were in close agreement with those obtained from using the FFT version.

The first-difference prewhitened true Dryden spectrum is shown in figure 16; it is the product of the first-difference transfer function squared $|H(f)|^2$ and the true Dryden spectrum of equation (6). Since this curve is concave upward in the frequency range of the first several points of the spectrum, the estimates at these

points should contain positive bias errors arising from the window convolution described earlier. The postdarkener function (eq. (3)) in this frequency range is quite large, going to infinity at zero frequency; therefore, the postdarkening operation could transform the positive bias errors mentioned earlier into a curve similar to figure 13 or to figures 14 and 15.

To determine the precise amount of bias error caused by window convolution alone, a digital simulation of the Blackman-Tukey window convolution was performed as follows: The Hann spectral window,

where $\Delta f = \frac{1}{P}$ is the same as the Δf for figure 13 (see fig. 2(b)),

was represented on the interval $-8\ \Delta f$ to $8\ \Delta f$ as a 16 001-point sampled-data function. The prewhitened Dryden spectrum, including the negative-frequency plane, was also represented as a sampled-data function, with the same frequency spacing $\frac{\Delta f}{1000}$. The convolution integral was then computed at frequencies $i\ \Delta f$, where $i = 1, 2, \ldots, 10$ from the vector dot product of the shifted window and the theoretical curve. These 10 values were then postdarkened. The maximum bias occurred at $i = 1$, the spectrum at this point being approximately 1.5 times the correct value. Therefore, window convolution bias alone is not sufficient to account for the Blackman-Tukey results of figure 13, although it probably accounts for a substantial portion of the error in the two direct-transform algorithm results (figs. 14 and 15).

Two additional possible sources of error are now discussed. The autocorrelation function (see ref. 8, p. 13) indicates a high correlation between successive values of the time function (that is, at $\tau = \Delta t$). This implies that the application of the first-difference filter (eq. (2)), where the subtraction of successive points occurs, results in a loss of precision of the filter output relative to the precision of the input. (Points of nearly equal

magnitude are being subtracted.) Because such round-off error is random, it adds a certain amount of white noise to the prewhitened spectrum and thus contributes to the positive bias near zero frequency. However, figures 14 and 15 suggest that this contribution is evidently small. Such an effect would be equal in all three algorithms.

The final source of error considered here is the computer round-off error which occurs during calculation of the power spectral density from the time function. It should be noted that the Blackman-Tukey algorithm requires more than twice the number of computations as the other two algorithms (two full-length transforms plus one short transform for Blackman-Tukey, as compared with one full-length transform for the other two algorithms). The lag-product Blackman-Tukey algorithm requires still more calculations than the FFT version does. This computer round-off error is also random and could contribute a flat error spectrum to the positive bias near zero frequency. Therefore, if the white noise spectrum for the Blackman-Tukey algorithm is greater than those for the other two algorithms, and is as large as the approximation in figure 16, it could contribute to window convolution bias error and could account for the effect observed in figure 13. This is probably the source of the large difference observed in figure 13 in comparison with the differences shown in figures 14 and 15.

The results obtained when a very wideband window is employed with the Blackman-Tukey algorithm $\left(B_e = 0.391 \text{ Hz} \quad \text{or} \quad B_{e_{k'}} = 2.56\right)$ are now examined. Windows of about this size, together with first-difference prewhitening, were generally used for processing turbulence data before the FFT was available. (The data presented here, although processed by the use of FFT, are mathematically equivalent to the previously used lag-product Blackman-Tukey method, as indicated earlier.) The spectra with and without prewhitening are shown in figures 17(a) and 17(b), respectively. The corresponding pair for an intermediate bandwidth of 0.195 Hz $\left(B_{e_{k'}} = 1.28\right)$ are shown in figure 18. In the past, the lowest frequency estimate from the

Blackman-Tukey lag-product algorithm was assumed to be at exactly zero frequency and was discarded because the postdarkener filter function (eq. (3)) goes to infinity at zero frequency, and also because the turbulence measurements were known to be contaminated by instrumentation drift problems at frequencies close to zero.

If this lowest frequency point (which is postdarkened and displayed at $\frac{\Delta f}{4}$ on these plots, as indicated earlier) is discarded, the lowest frequency estimate then appears at 0.195 Hz in figure 17(a) and agrees very well with the reference curve. The next power estimate, which appears at 0.39 Hz, slightly underestimates the reference curve. For the non-prewhitened result (fig. 17(b)), the corresponding points underestimate and overestimate the true spectrum to a somewhat greater extent, so that for this large bandwidth the prewhitened results are indeed better (when the "zero frequency" estimate is ignored). On the other hand, figure 18 shows that when the bandwidth is decreased to 0.195 Hz, the non-prewhitened results are better, whether the lowest frequency power estimate is ignored or not.

These intermediate bandwidth results indicate that in the past, the use of first-difference prewhitening could have obscured a possible knee in the experimental data. (See fig. 18(a).) It should be recalled, however, that interest at the time was directed toward defining the higher frequency end of the spectrum (using wider bandwidths) with relatively short turbulence sample lengths.

As an illustration of the minimum sample length required for this bandwidth, assume that use of 30 d.o.f. in equation (7) does not result in intolerably large random power fluctuations. A B_{e_k}, of 1.28 produces a lowest power estimate at about the knee peak location, and requires a sample length of only about 77 sec for the ratio of $\frac{L}{V}$ assumed here.

Trade-off factors for real data are such that analyzing power spectra to longer wavelengths (and thus lower frequencies) requires a smaller B_e $\left(\text{also } B_{e_k'}\right)$, which in turn requires longer sample lengths to suppress the random statistical fluctuations of the power estimates. As $B_{e_k'}$ becomes smaller than about 1.2, which it must to analyze longer wavelengths or adequately resolve a possible knee, first-difference prewhitening is no longer desirable and, in fact, increases the bias error if the true spectrum is similar in shape to the Von Kármán or Dryden turbulence model.

The practical procedure should therefore consist of choosing a minimum B_e based on the sample length available. A determination or "rule of thumb" in terms of $B_{e_k'}$ will not suffice, since it must be assumed that L, the integral scale of turbulence, is not known at this point, nor can it be assumed that the Von Kármán or Dryden turbulence model is applicable, for that matter. Obtaining sufficiently long data samples to suppress the random error, which is considerably greater than bias error at this point, poses a problem. It has been suggested that the ensemble averaging algorithm could be used to alleviate the difficulty by combining segments of data obtained from repeated passes through the same turbulent area, or from different turbulent areas known to have been generated by the same meteorological process. Such a procedure might be questionable, however, since it would be based on the assumption that the spectral characteristics of all the segments would be identical.

If it is necessary to use an intermediate size bandwidth (that is, around 0.1 to 0.2 Hz) because of sample length limitations, some special type of prewhitening other than first difference could be beneficial. The spectral shape and frequency location of the knee must then be estimated or known in advance. For example, long samples in a specific meteorological condition could have previously been processed and would lend confidence to a particular shape and knee location. Care must be exercised, however, since a "wrong guess" would lead to greater distortion of the final result, rather than less.

If wide bandwidth windowing is necessary (that is, B_{e_k}, of about 2.6 and larger) because of limitations in sample length, then first-difference prewhitening is required for accurate spectral estimates to be obtained, as shown by the results given in figure 17. It is also recommended that the "zero frequency" estimate obtained with the Blackman-Tukey algorithm be discarded, as was done in the past, especially in view of the large distortion obtained near zero with this algorithm when using first-difference prewhitening. The spectral knee could not, of course, be resolved in this case since the lowest frequency spectral estimate obtained would be above the knee location.

CONCLUDING REMARKS

A study was made of the effects of prewhitening on determination of power spectra of atmospheric turbulence at long wavelengths. A synthetic time history was generated by combining sine waves of random phase and frequency, amplitudes being adjusted to produce a power spectrum of known shape approximating that of atmospheric turbulence. The synthetic time history was then used in a computational experiment to assess bias errors in power spectra computed with and without prewhitening.

Results of this experiment show that for minimum bandwidths deemed practical for processing 10-minute data samples (equivalent spectral bandwidth of 0.025 Hz), the finite bandwidth bias errors, for each of three power spectral density algorithms implemented by the fast Fourier transform, are negligible and are a great deal smaller than the random type errors expected. Prewhitening is therefore not recommended when power spectral estimates are obtained by employing these narrow bandwidths. First-difference prewhitening in particular was shown to have an undesirable effect upon power estimates at the lowest frequencies, especially when used with the Blackman-Tukey algorithm.

First-difference prewhitening is not recommended when intermediate size bandwidths (that is, around 0.1 to 0.2 Hz) are used, although some special type of prewhitening could be beneficial if the spectrum shape and knee location were adequately known. First-difference prewhitening is, however, recommended for power estimates above the "knee" frequency for bandwidths of 0.3 Hz and greater.

For model studies of spectral bias errors of a random time series, the function-generation technique used here is favored over the technique used by Otnes, Nathans, and Enochson (AFFDL-TR-69-11). The random spectral errors observed can be made arbitrarily small for a given function length by summing a sufficiently large number of random-frequency, random-phase sinusoids.

Langley Research Center
National Aeronautics and Space Administration
Hampton, VA 23665
August 11, 1976

APPENDIX

SYNTHETIC RANDOM TIME FUNCTIONS

Synthetic random time functions are often used in lieu of random time functions representing physical random processes. For example, aircraft flight simulators include responses to atmospheric turbulence velocities that are approximated by random time functions generated by suitably filtered electronic noise signals or by digital computer programs. This approach has the advantage of providing (a) arbitrarily chosen intensities, (b) stationarity if desired, and (c) simulation of the effect of various integral scales of turbulence.

In the construction of digitally generated random time functions, care must be taken to assure that the significant features of the physical random process are modeled adequately. The digital function, of course, must be constructed from a finite number of randomly chosen elements. There is the possibility that a particular choice of elements may not satisfactorily model the physical process for certain purposes even though the probability distribution and spectral shape are correct and the random time function appears to be representative of the physical process. This appendix presents an example of such an approximation (designated Signal B) for atmospheric turbulence with respect to the assessment of effects of various spectral windows. An alternative approximation, described in the text of this report (designated Signal A) is also discussed, with the generation technique described in more detail.

Signal B

Signal B was generated as follows, in accordance with reference 3:

APPENDIX

(1) 16 384 random numbers were generated. These numbers were independent, Gaussian, with zero mean and unit variance. The Δt was selected to be 0.04 sec.

(2) The complex FFT of these data was multiplied by the desired transfer function $H(f)$, the square root of equation (6), with $L = 191$ m (627 ft) and $V = 183$ m/s (600 ft/sec).

(3) The inverse FFT was then taken.

The resulting function from step (2) is defined only at harmonically related frequencies $n\,\Delta f$ where $n = 0, 1, 2, \ldots,$ 8192, and where $\Delta f = \dfrac{1}{16\ 384\ \Delta t}$. The inverse transform of each of these is a pure harmonic sinusoid with frequency $n\,\Delta f$; therefore Signal B is the sum of the harmonic sinusoids. The spectrum of the model differs significantly from that for atmospheric turbulence since the turbulence spectrum is continuous and contains nonharmonically related components.

The analysis of Signal B by the boxcar spectral window (fig. 1(b)) in equation (1) yields no bias error at all as a result of the coincidence of nulls in the spectral window with the spectral values of Signal B at $f = \dfrac{n}{P}$.

A secondary indication of the improper behavior of Signal B in comparison to atmospheric turbulence can be seen in figure 19. Here Signal B was processed by the frequency averaging algorithm $M = 16$ with the Hann window (fig. 19(a)), and with the boxcar window (fig. 19(b)). It is apparent that the boxcar window suppressed random error to a considerably greater extent than did the Hann window, an effect exactly opposite from that which would be obtained with real turbulence data. A probable explanation is that the Hann window does not have nulls at $\dfrac{-1}{P}$ and $\dfrac{+1}{P}$ as does

the boxcar window. (See fig. 2(b).) The inherent characteristics of Signal B would thus appear to make it unsuitable for use in the present study.

Signal A

Signal A was also generated to contain 16 384 data points spaced at an interval of $\Delta t = 0.04$ sec, which yielded a Nyquist frequency 12.5 Hz. The signal was constructed as follows: The total frequency range, 0 to 12.5 Hz, was partitioned into 1250 equal intervals of 0.01 Hz each. For each 0.01 Hz interval, a uniform-distribution random number generator was used to select a frequency f and a phase angle θ, with f lying in the interval and $-\pi \leq \theta \leq \pi$. For each frequency and phase angle thus selected, the sinusoid $H(f) \sin (2\pi f n \Delta t + \theta)$ (where $n = 1, 2, \ldots, 16\,384$) was computed. Signal A is the sum of these 1250 sinusoids. The transfer function $H(f)$ used to adjust the amplitude of the individual sinusoids was, of course, the same one used for Signal B, that is, the square root of equation (6), with $L = 191$ m (627 ft) and $V = 183$ m/s (600 ft/sec).

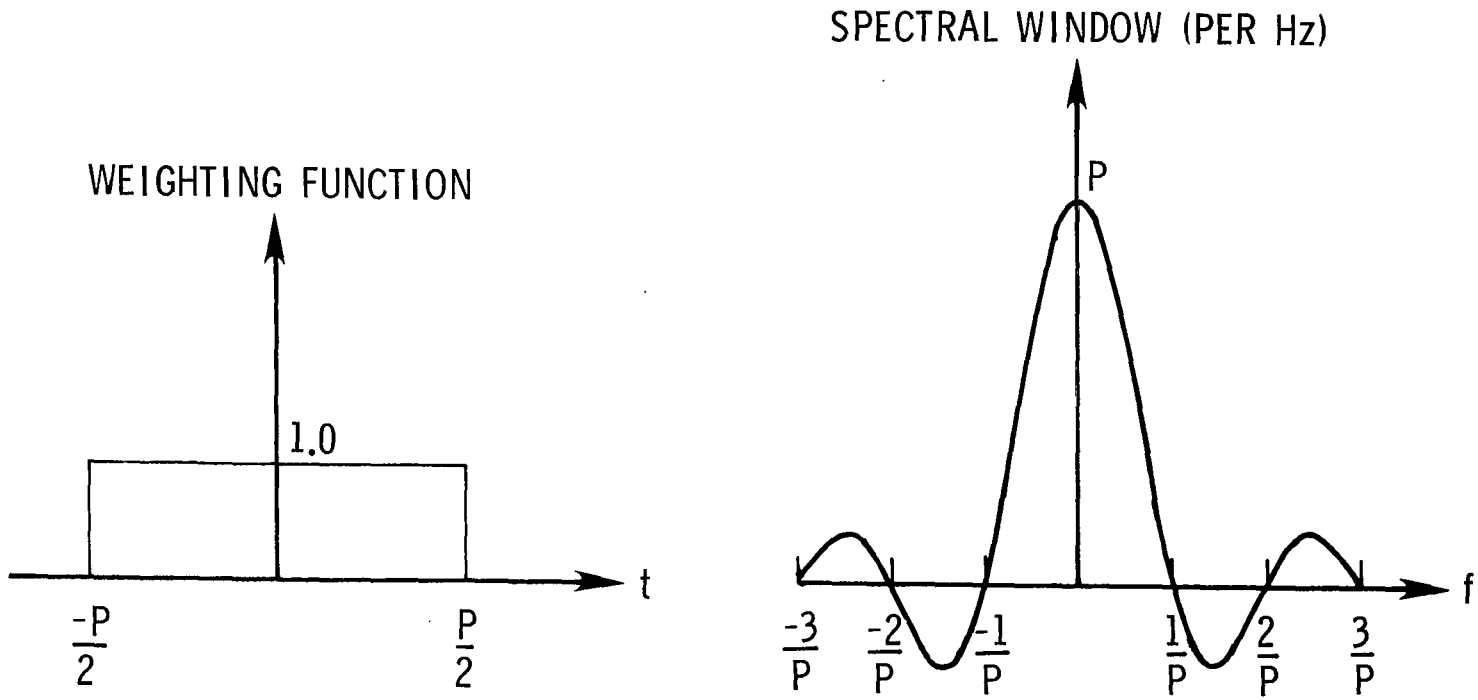
Since both the frequency and phase are made random by the construction procedure, Signal A is not made up exclusively of harmonic components which coincide with boxcar window nulls. The approximation of atmospheric turbulence is thus improved. As described in the main body of this report, a short segment of Signal A is shown in figure 8. The cumulative distribution function, shown in figure 9, is very nearly Gaussian, as desired, and as expected from the procedure used.

An additional point worth noting is that the frequency spacing of the sinusoids, one in each 0.01 Hz interval, is such that the smallest bandwidth B_e employed to process Signal A (about 0.024 Hz) encompassed at least two sinusoids. Although not tested experimentally, the random power fluctuations obtained in the spectral analysis (for a given B_e) could in all likelihood be reduced

to an arbitrarily low level by increasing the number of sinusoids in each frequency interval. The effect would be similar to that achieved by increasing the sample length of actual turbulence data.

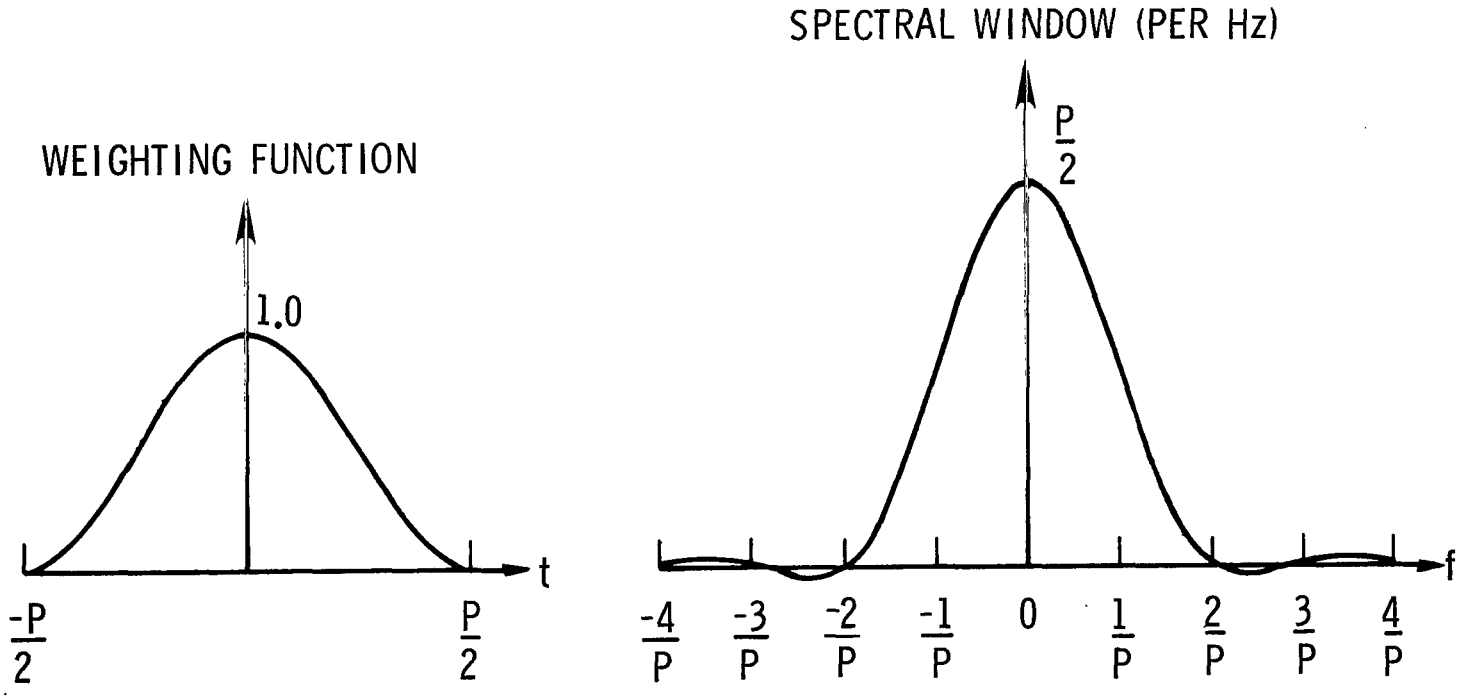
REFERENCES

1. Houbolt, John C.: Atmospheric Turbulence. AIAA J., vol. II, no. 4, Apr. 1973, pp. 421-437.
2. Press, Harry; and Tukey, John W.: Power Spectral Methods of Analysis and Their Application to Problems in Airplane Dynamics. Vol. IV of AGARD Flight Test Manual, Pt. IVC, Enoch J. Durbin, ed., Pergamon Press, 1962, pp. IVC:1 - IVC:41.
3. Otnes, R. K.; Nathans, H. A.; and Enochson, L. V.: A Procedure for Computing Power Spectral Density of Gust Data. AFFDL-TR-69-11, U.S. Air Force, May 1969.
4. Enochson, Loren D.; and Otnes, Robert K.: Programming and Analysis for Digital Time Series Data. SVM-3 (Contract NO014-67-C-0073), Shock & Vibration Information Center, U.S. Dept. Defense, 1968. (Available from DDC as AD 692 735.)
5. Welch, Peter D.: The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. IEEE Trans. Audio and Electroacoust., Vol. AU-15, no. 2, June 1967, pp. 70-73.
6. Brown, Thomas J.; Brown, Christine G.; and Hardin, Jay C.: Program for the Analysis of Time Series. NASA TM X-2988, 1974.
7. Otnes, Robert K.; and Enochson, Loren: Digital Time Series Analysis. John Wiley & Sons, Inc., c.1972.
8. Houbolt, John C.; Steiner, Roy; and Pratt, Kermit G.: Dynamic Response of Airplanes to Atmospheric Turbulence Including Flight Data on Input and Response. NASA TR R-199, 1964.
9. Airplane Strength and Rigidity - Flight Loads. Mil. Specif. MIL-A-008861A(USAF), Mar. 31, 1971.
10. Ward, Robert C.: Dynamic Data Analysis Techniques Used in the Langley Time Series Analysis Computer Program. NASA TM X-2160, 1971.



(a) Time. (b) Frequency.

Figure 1.- Boxcar window.



(a) Time.

(b) Frequency.

Figure 2.- Hann window.

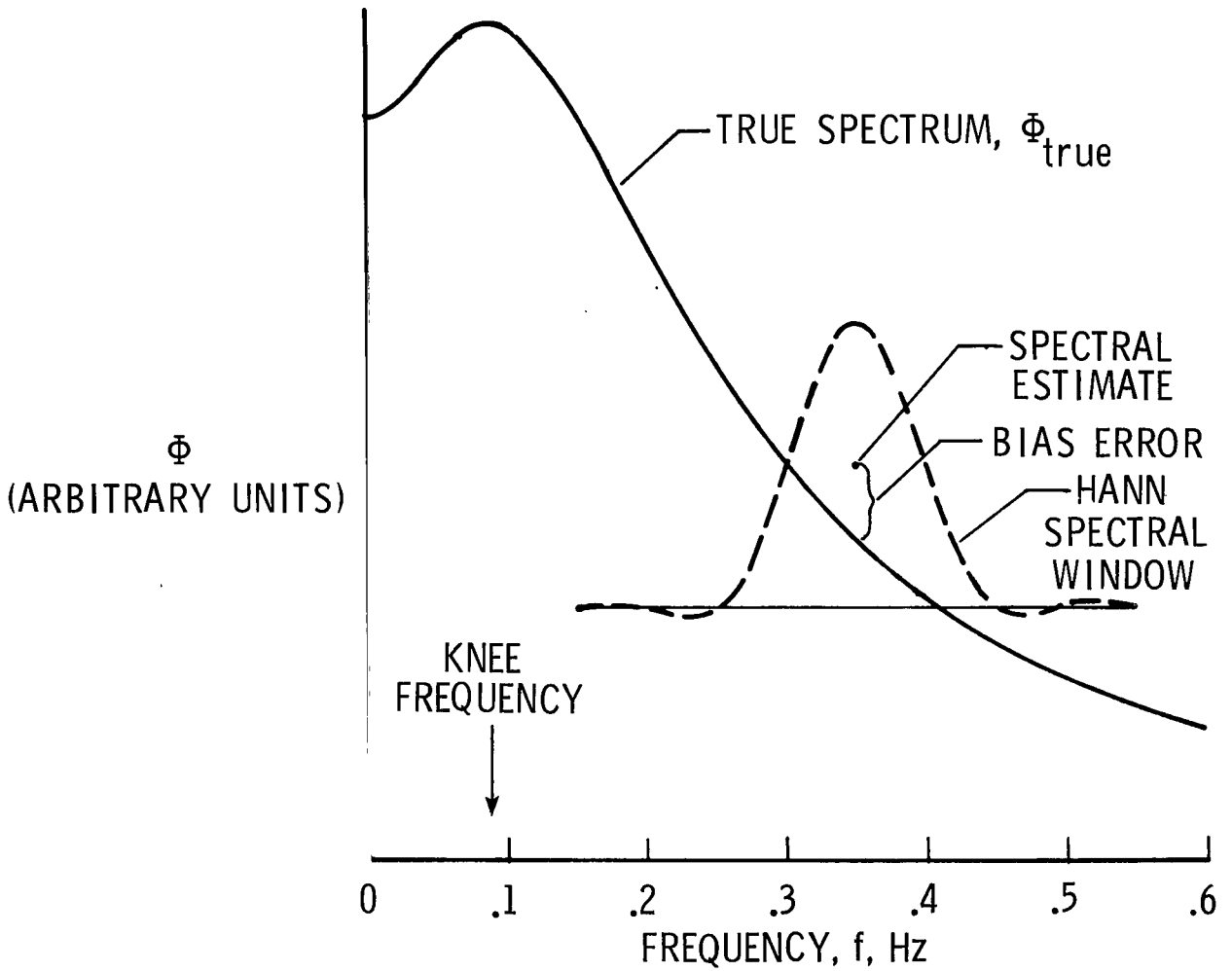
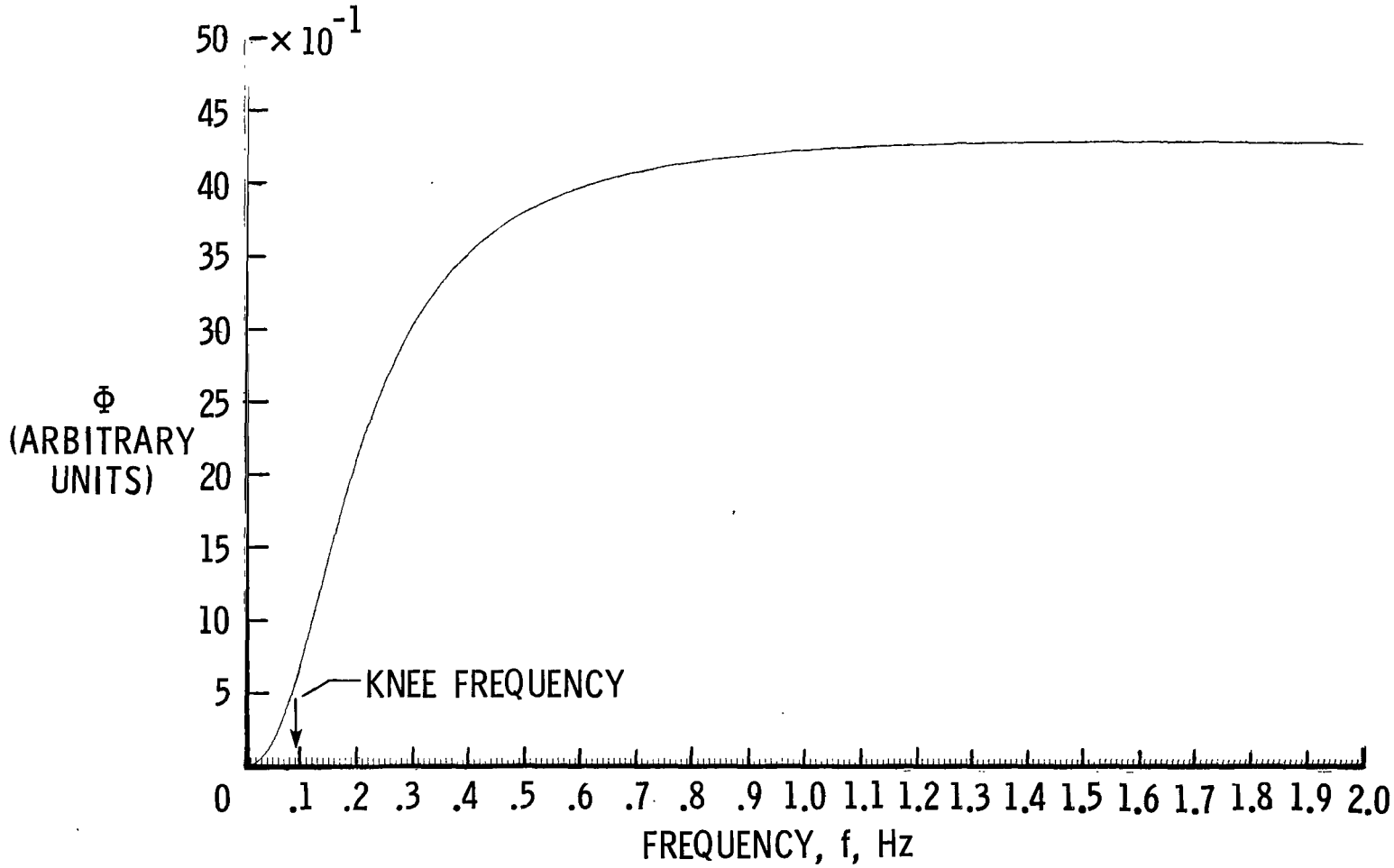
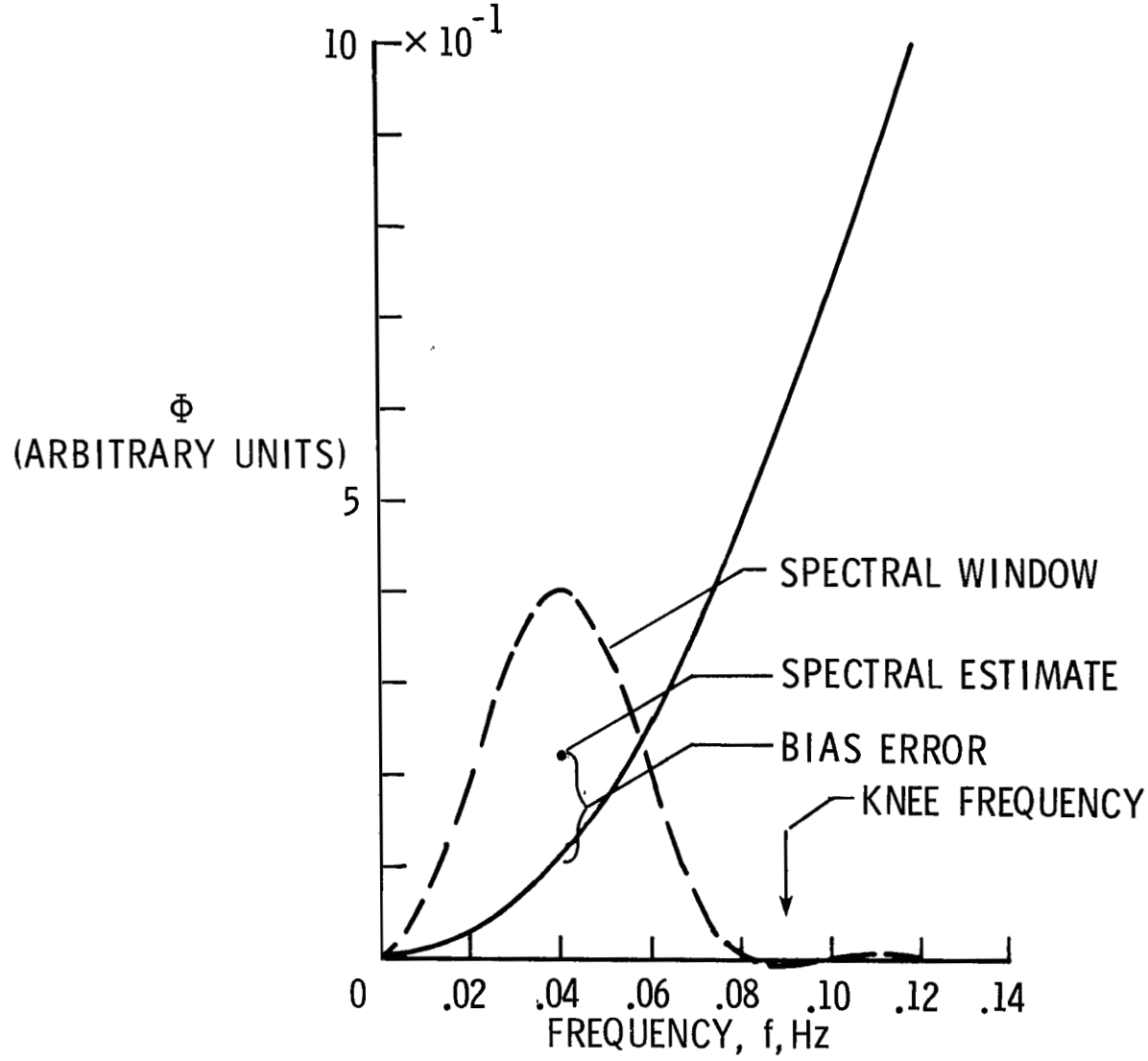


Figure 3.- Illustration of bias error resulting from convolution of spectral window with true spectrum.



(a) Frequency range, 0 to 2 Hz.

Figure 4.- True spectrum after applying first-difference prewhitening filter.



(b) Frequency range, 0 to 0.14 Hz.

Figure 4.- Concluded.

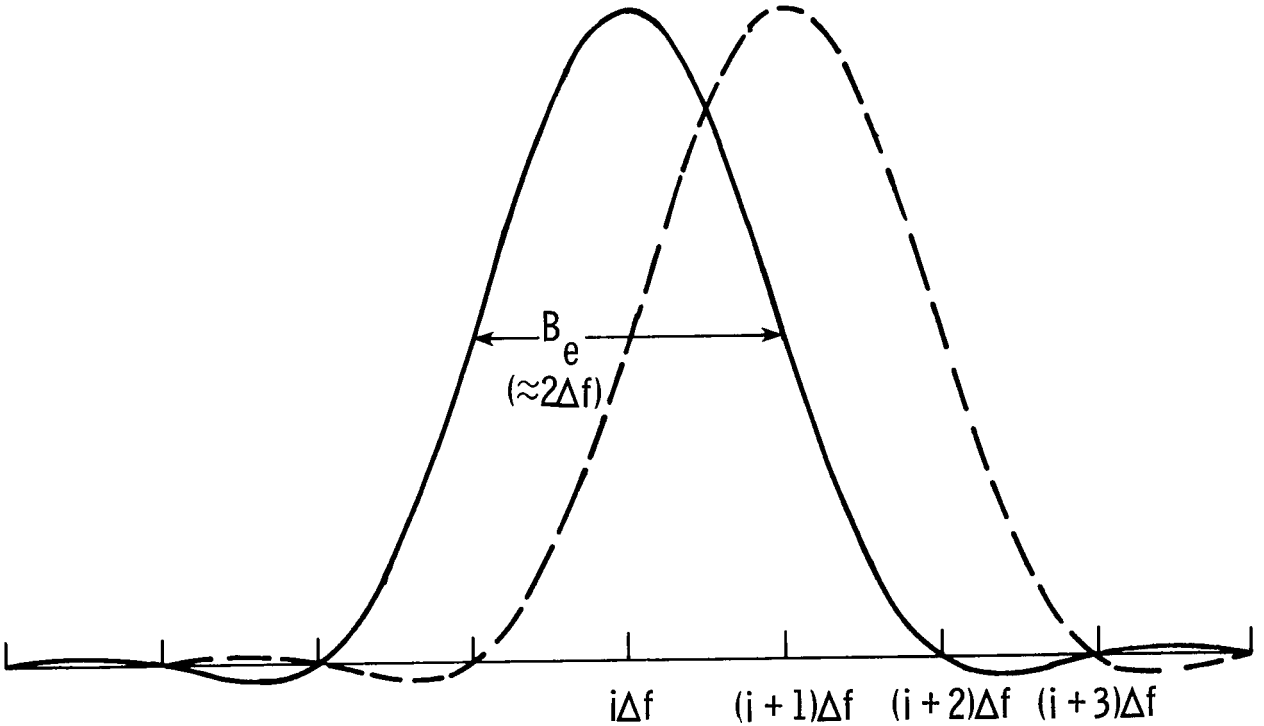


Figure 5.- Spectral bandwidth and window overlap for two adjacent spectral estimates (Blackman-Tukey algorithm, Hann window).

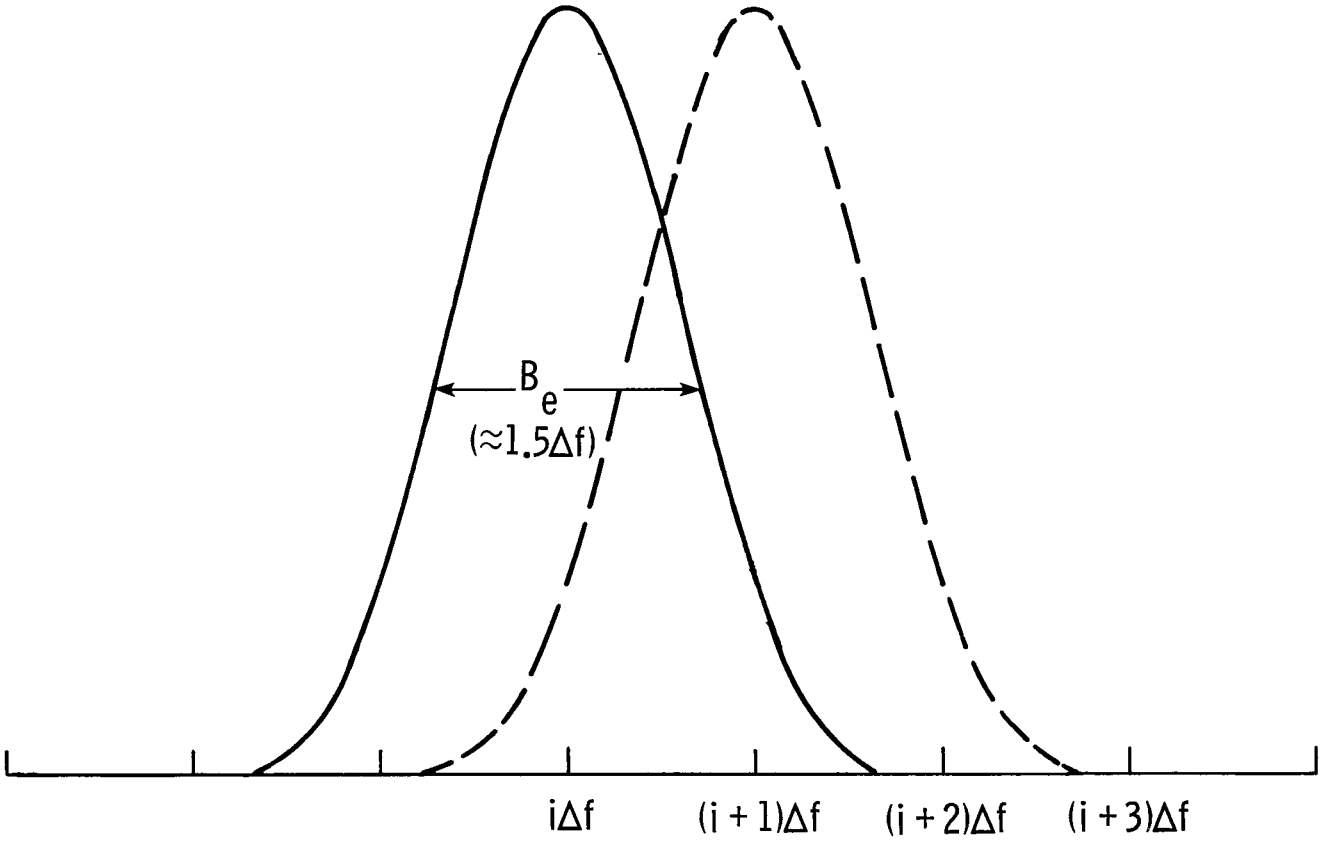


Figure 6.- Ensemble averaging window (squared Hann window) showing bandwidth and window overlap for two adjacent spectral estimates.

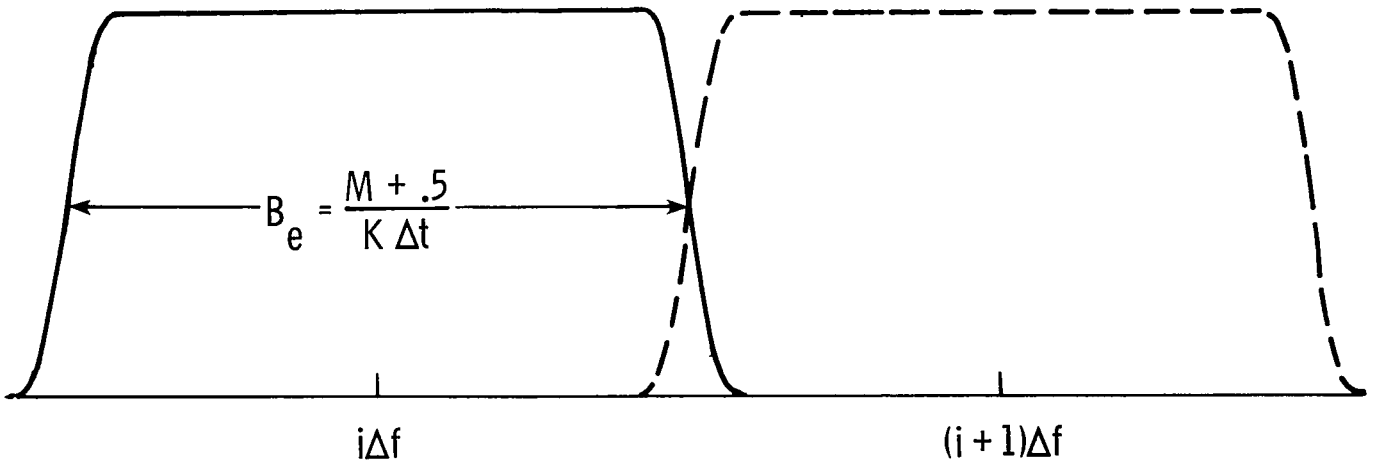


Figure 7.- Frequency averaging window for $M = 16$ (sum of 16 squared Hann windows), showing bandwidth and window overlap for two adjacent spectral estimates. M, number of contiguous raw power spectral density estimates used to compute each final estimate at $i\,\Delta f$; and K, total number of data points transformed.

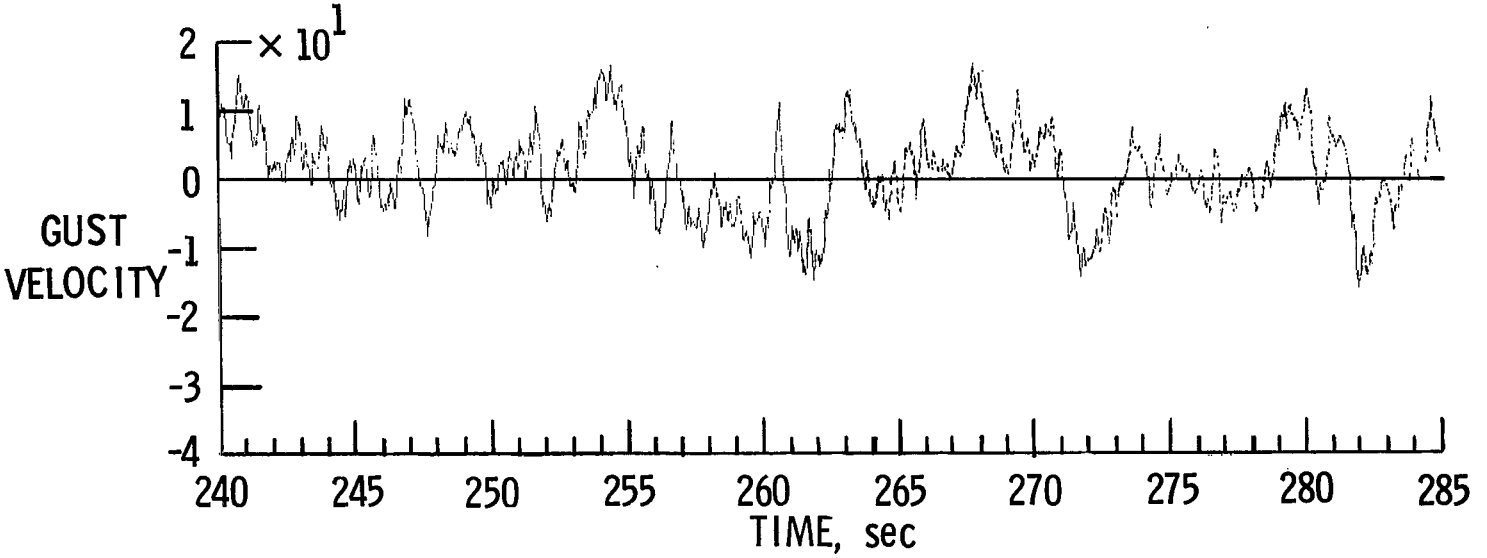


Figure 8.- Sample of artificially generated time history. Signal A.

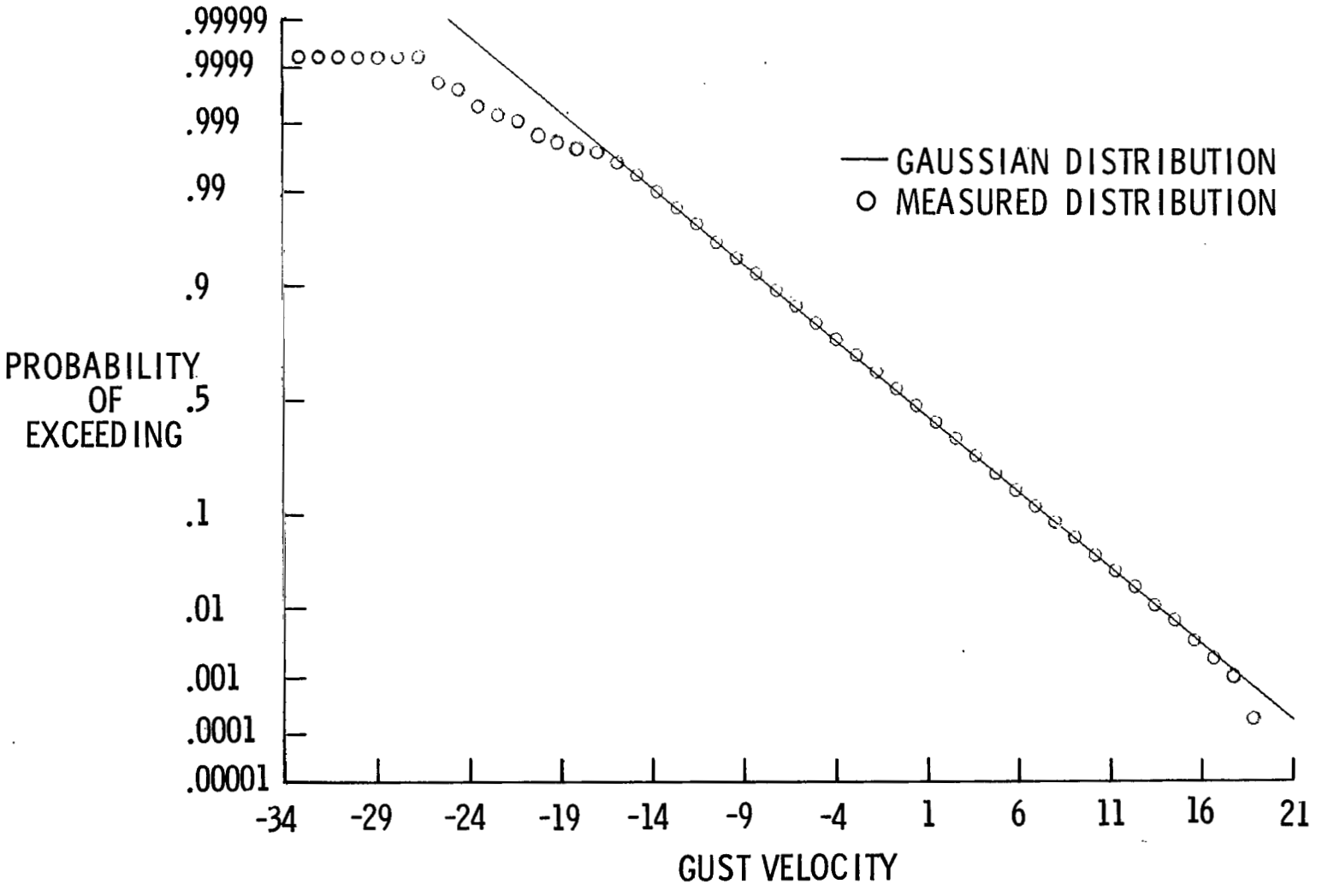


Figure 9.- Cumulative probability distribution of artificially generated time history. Signal A.

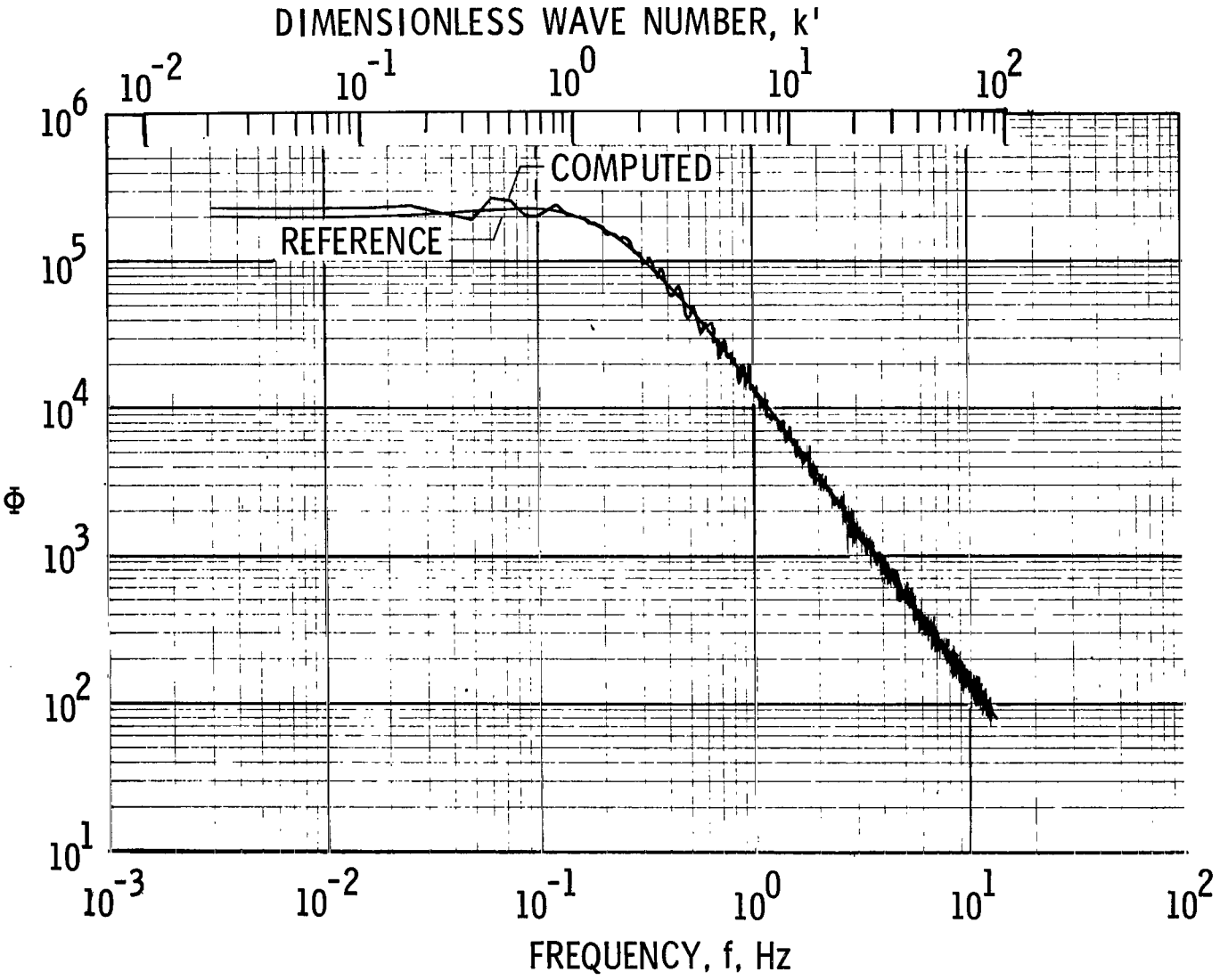


Figure 10.- Power spectrum obtained from Blackman-Tukey algorithm. No prewhitening; ℓ = 1025; Δf = 0.0122; B_e = 0.0244; and $B_{e_{k'}}$ = 0.160.

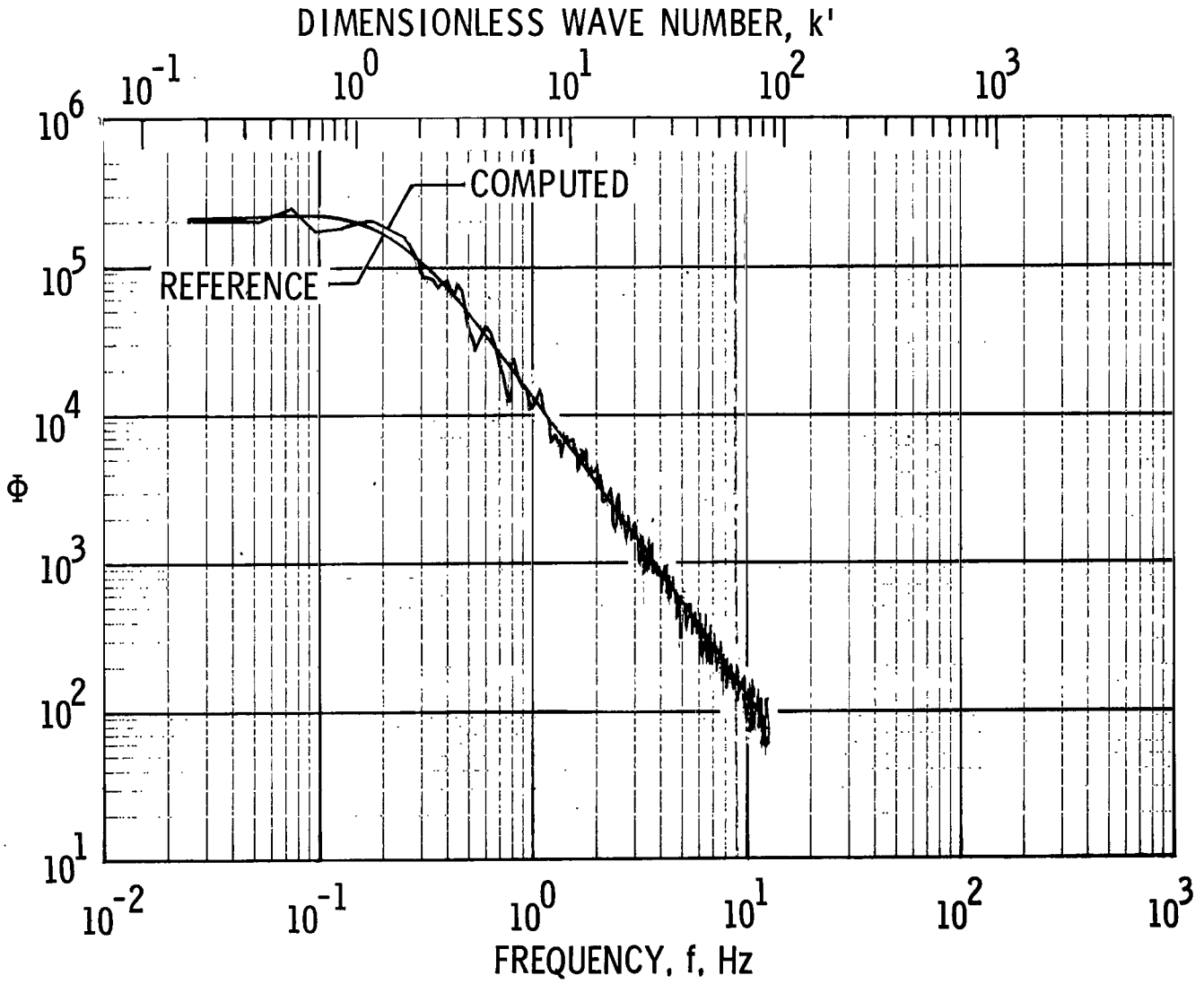


Figure 11.- Power spectrum obtained from ensemble averaging algorithm. No prewhitening; $\Delta f = 0.0244$; $B_e = 0.0366$; and $B_{e_{k'}} = 0.240$.

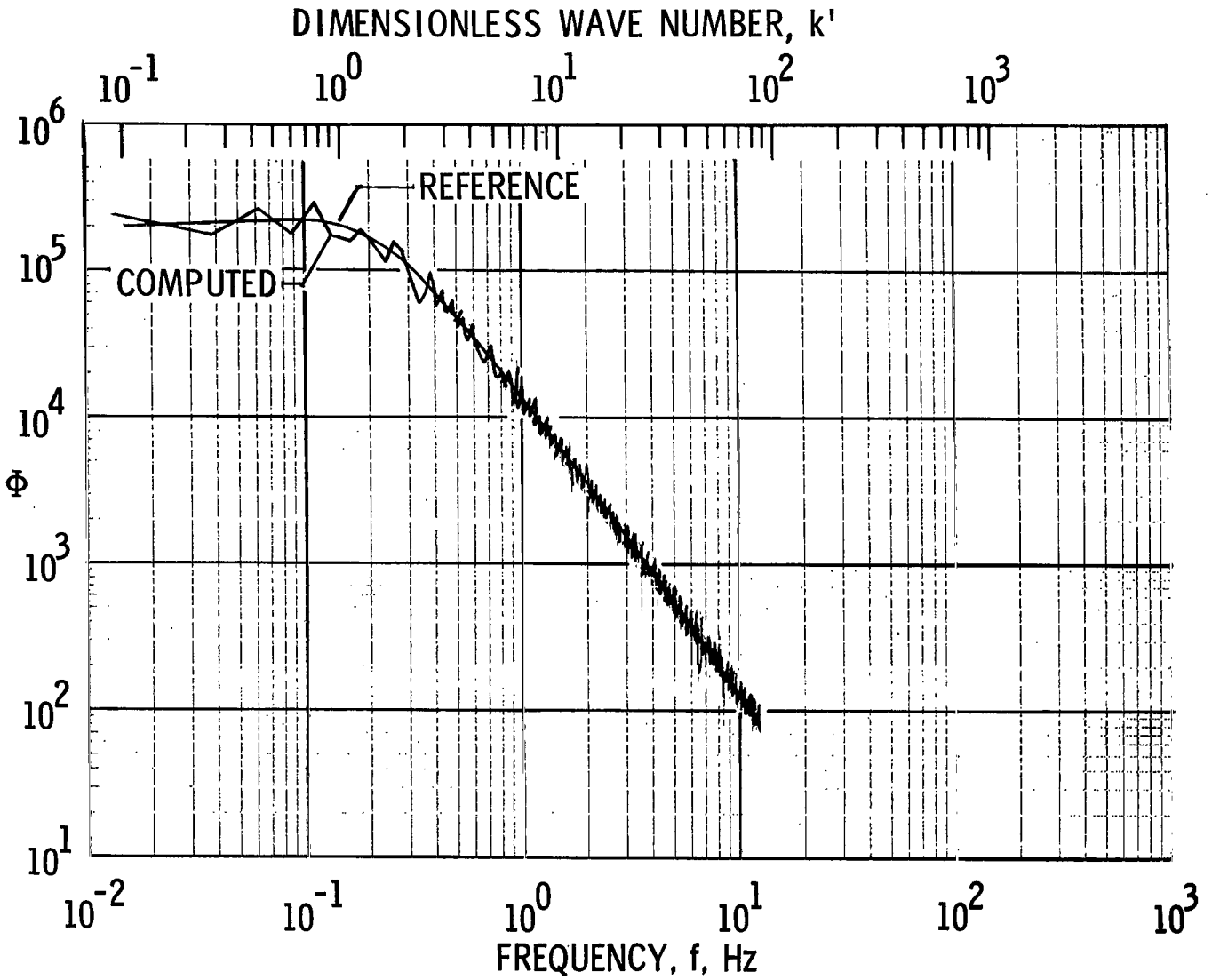


Figure 12.- Power spectrum obtained from frequency averaging algorithm. No pre-whitening; $\Delta f = 0.0244$; $B_e = 0.0260$; and $B_{e_{k'}} = 0.171$.

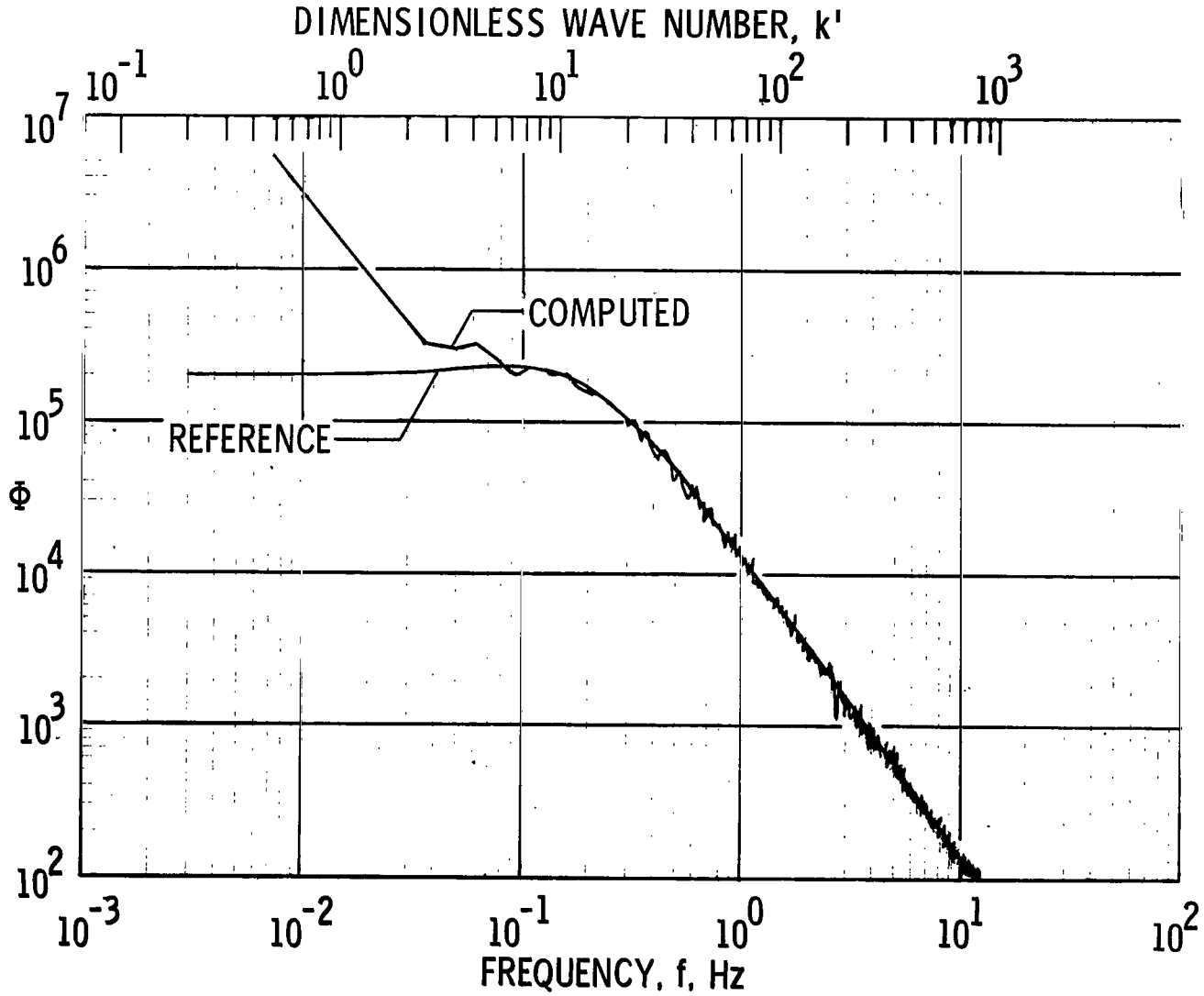


Figure 13.- Power spectrum obtained from Blackman-Tukey algorithm. First-difference prewhitening; ℓ = 1025; Δf = 0.0122; B_e = 0.0244; and $B_{e_{k'}}$ = 0.160.

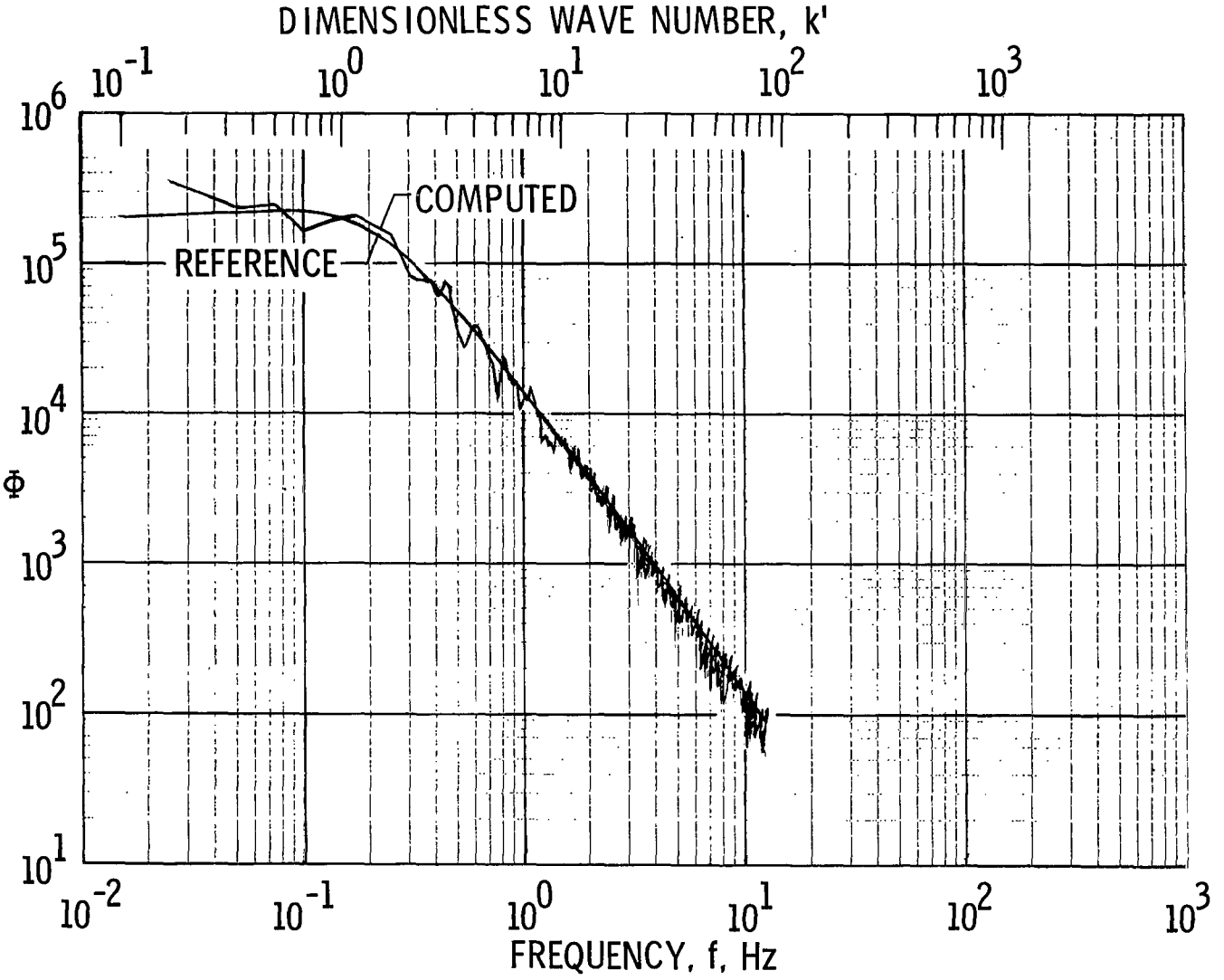


Figure 14.- Power spectrum obtained from ensemble averaging algorithm. First-difference prewhitening; $\Delta f = 0.0244$; $B_e = 0.0366$; and $B_{e_{k'}} = 0.240$.

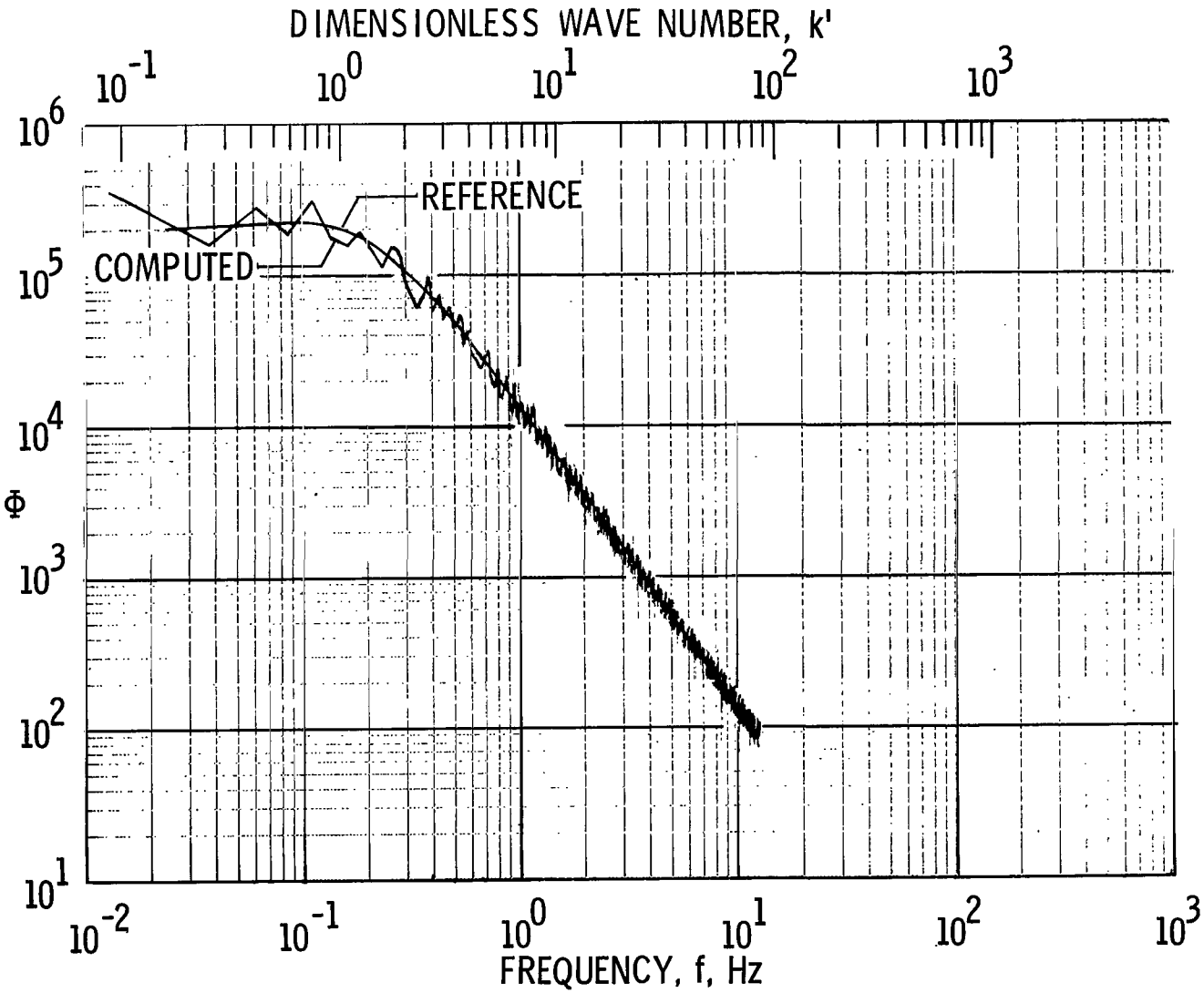


Figure 15.- Power spectrum obtained from frequency averaging algorithm. First-difference prewhitening; $\Delta f = 0.0244$; $B_e = 0.0260$; and $B_{e_{k'}} = 0.171$.

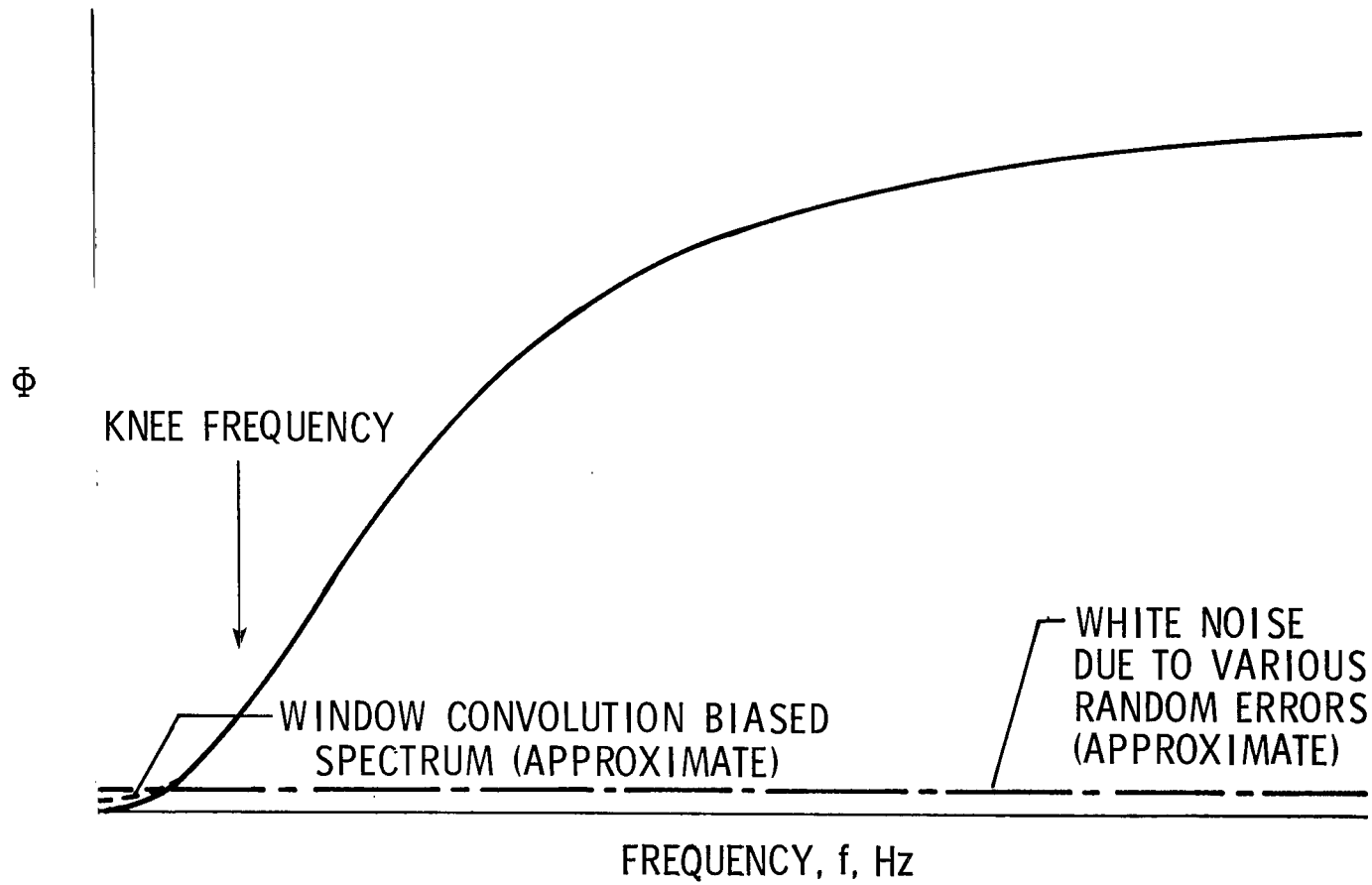
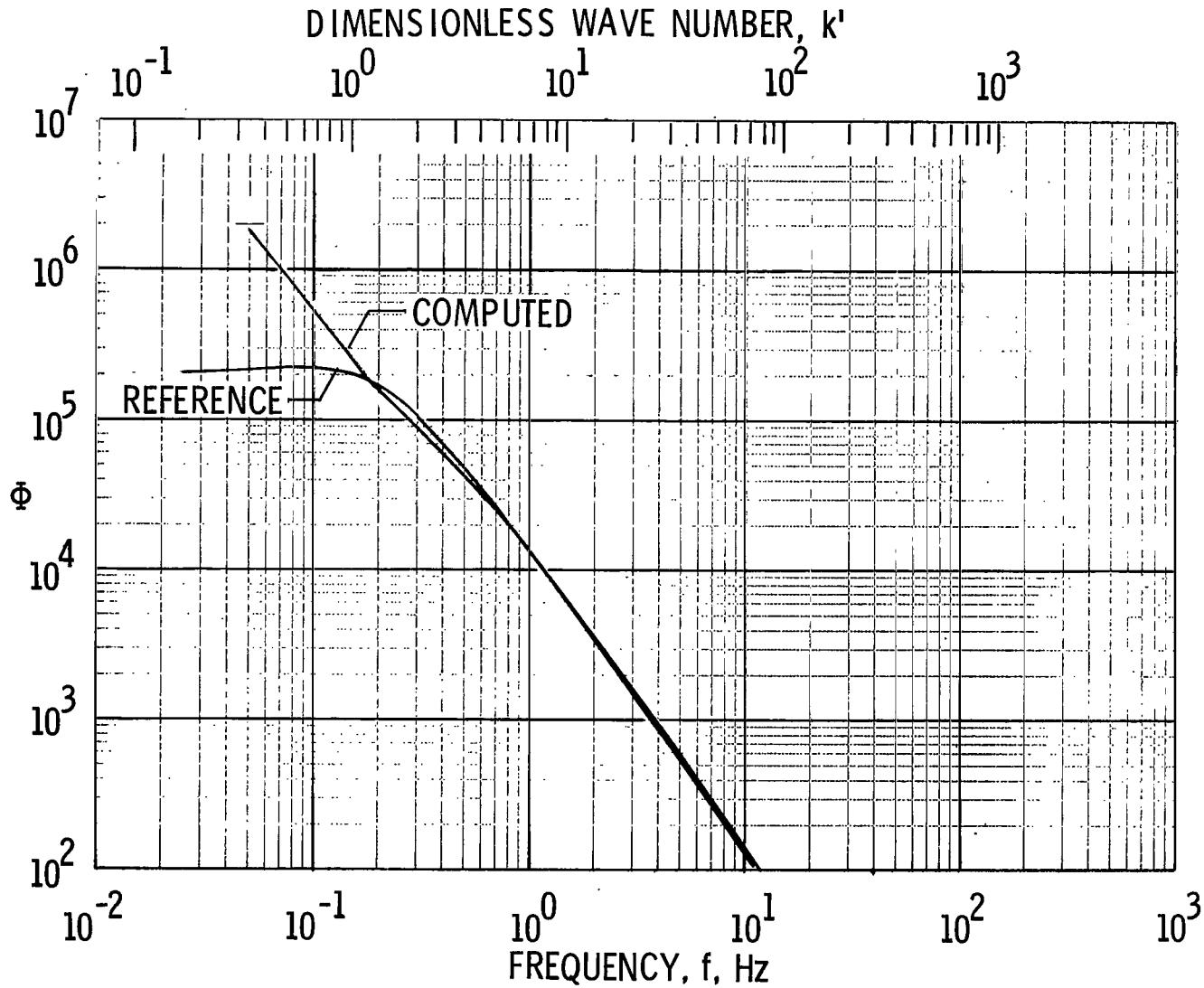
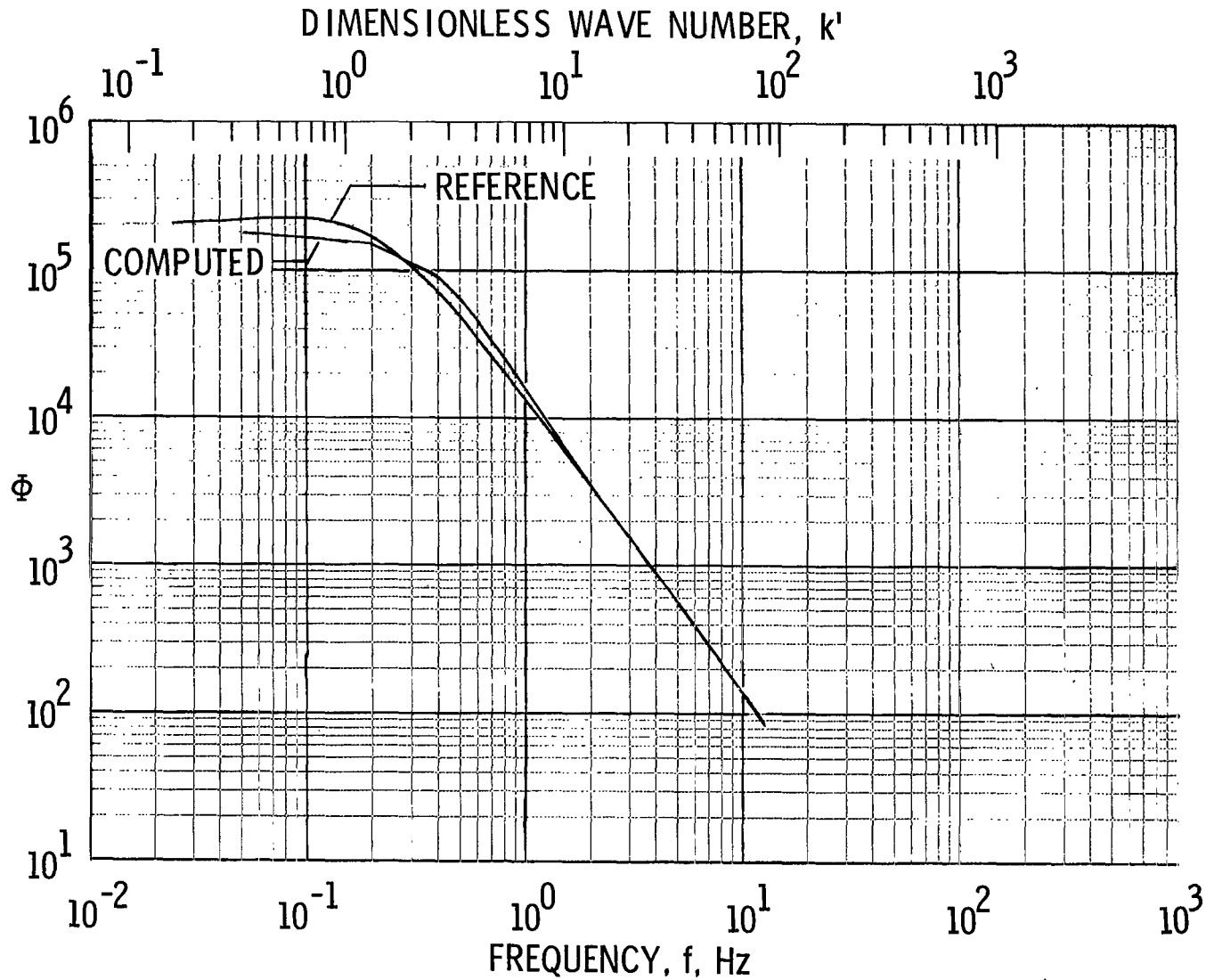


Figure 16.- First-difference prewhitened Dryden spectrum.



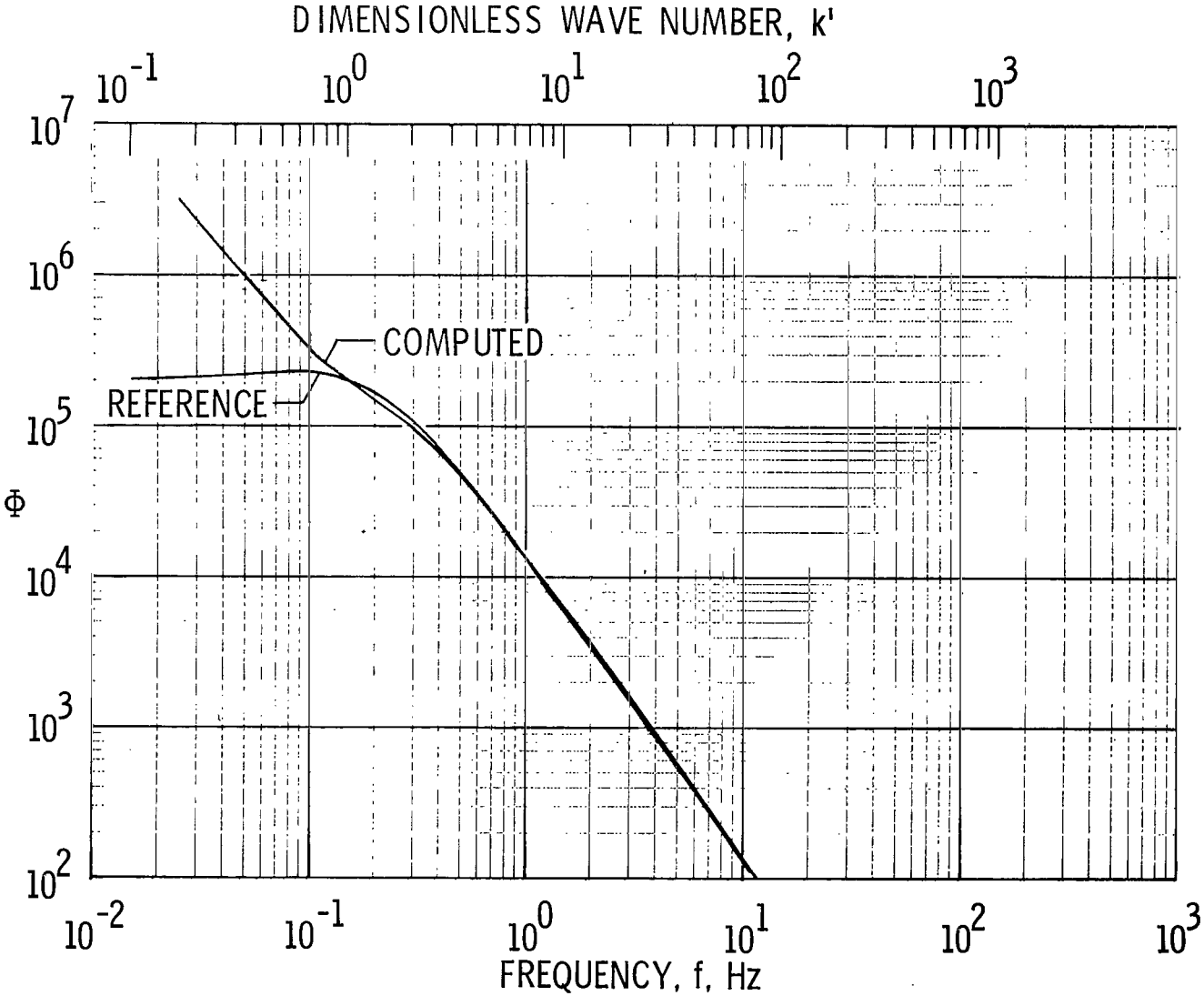
(a) Prewhitened.

Figure 17.- Comparison of prewhitened results and non-prewhitened results for wideband Blackman-Tukey algorithm. ℓ = 65; Δf = 0.195; B_e = 0.391; and $B_{e_{k'}}$ = 2.56.



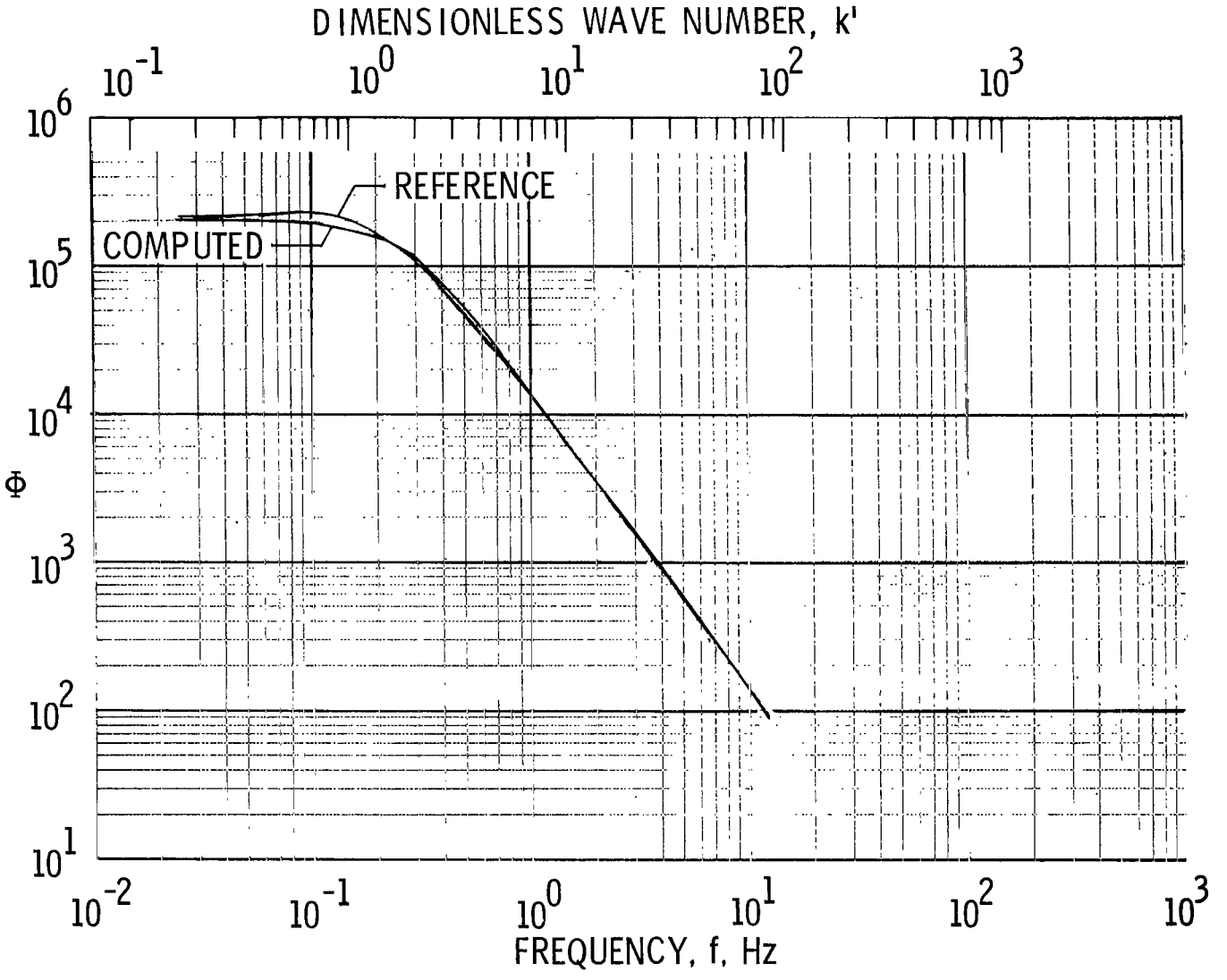
(b) Non-prewhitened.

Figure 17.- Concluded.



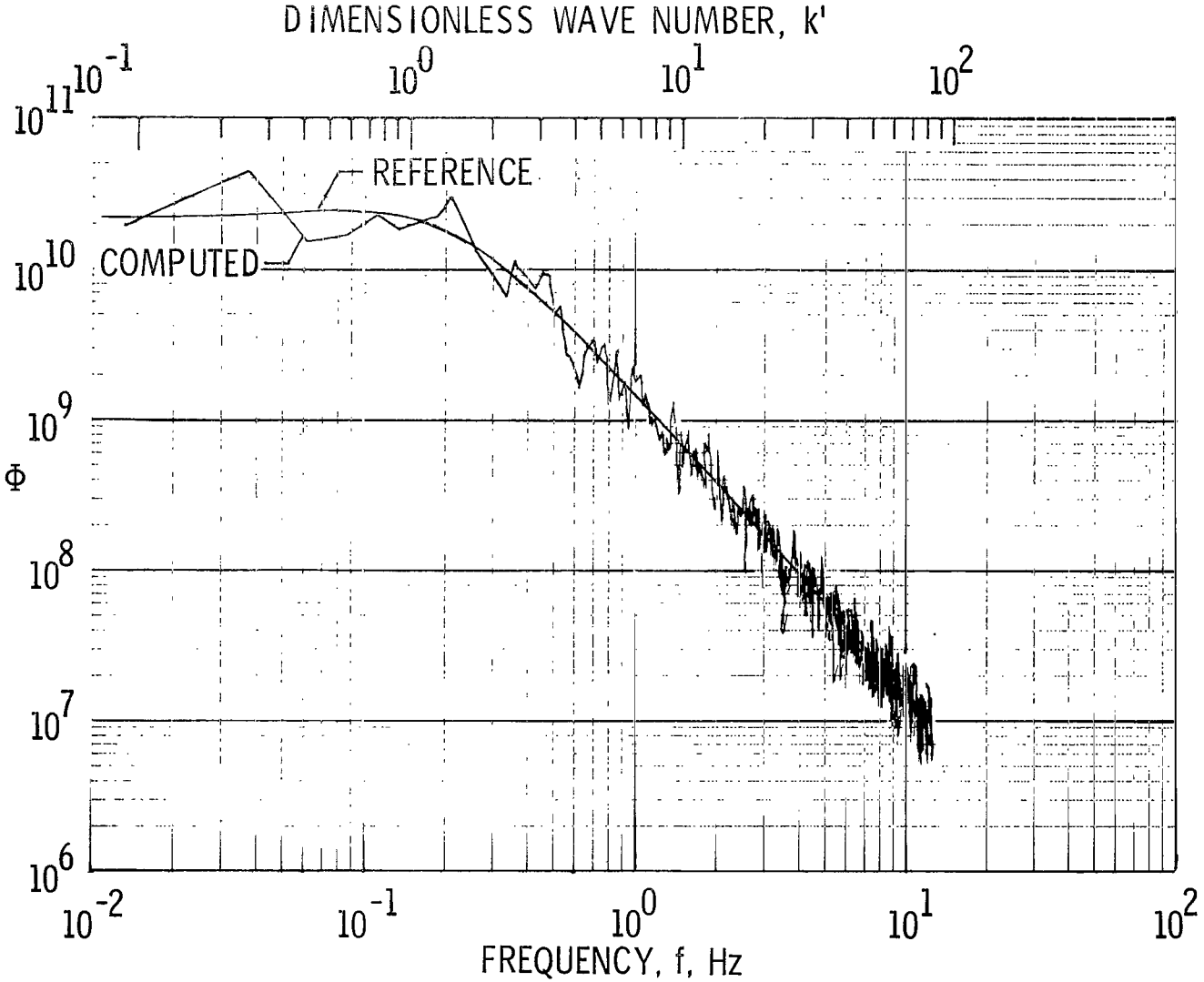
(a) Prewhitened.

Figure 18.- Comparison of prewhitened results and non-prewhitened results for wideband Blackman-Tukey algorithm. ℓ = 129; Δf = 0.0977; B_e = 0.195; and $B_{e_{k'}}$ = 1.28.



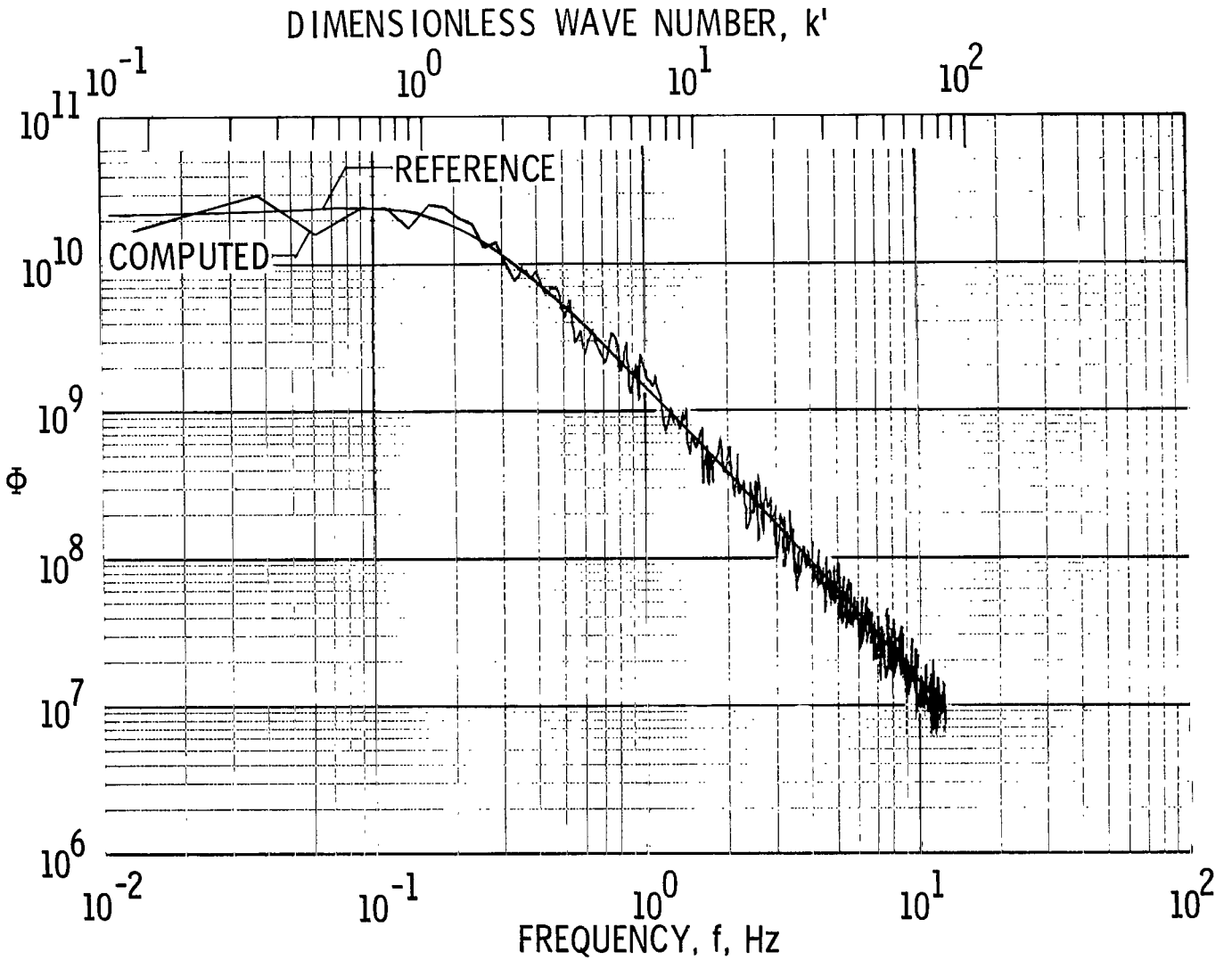
(b) Non-prewhitened.

Figure 18.- Concluded.



(a) Hann window.

Figure 19.- Comparison of results obtained with Hann window and results from boxcar window from one 16 384-point realization of Signal B frequency averaged over 16 power estimates (that is, M = 16).



(b) Boxcar window.

Figure 19.- Concluded.

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons. Also includes conference proceedings with either limited or unlimited distribution.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include final reports of major projects, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546