Sampling and Aliasing During the NORPAX Hawaii-to-Tahiti Shuttle Experiment

ERIC FIRING and ROGER LUKAS

Joint Institute for Marine and Atmospheric Research, University of Hawaii, Honolulu

We examine leakage and aliasing problems in a seasonal multiple linear regression model of data taken during the NORPAX Hawaii-to-Tahiti Shuttle Experiment. The model includes a mean, trend, annual harmonic, and semiannual harmonic. Inclusion of the trend can exacerbate leakage, particularly for record lengths of 13 months or less. This is not a major problem for the 15-month shuttle record. The irregularity of the shuttle shipboard time series reduces the potential aliasing of single unresolved high frequencies. Confidence limits for regression model coefficients are determined empirically from Monte Carlo subsampling experiments on moored current meter and sea level data taken during the shuttle. These limits are generally narrower than the standard regression error estimates because there is lack of fit in the regression model. Confidence limits used here pertain only to the determination of model coefficients for a given data set; they say nothing about whether the annual cycle during the shuttle period is typical of any other time interval. The shuttle sampling at each of three longitudes (150°, 153°, and 158°W) was in general not adequate to reveal significant differences in low-frequency signals from one longitude to another. If data from the three longitudes are analyzed together without regard for longitude, the multiple linear regression parameters for temperature and zonal velocity component are fairly well determined. The meridional velocity component is so dominated by higher-frequency variability that the model parameters do not differ significantly from zero.

1. INTRODUCTION

The NORPAX Hawaii-to-Tahiti Shuttle Experiment was a study of currents and hydrography in the central equatorial Pacific during one seasonal cycle [Wyrtki et al., 1981]. Measurements were made on 15 cruises at roughly monthly intervals. The central purpose of this experiment was to observe one realization of the seasonal cycle and to accurately determine the mean during that cycle. In this paper we quantify how well that was accomplished. We were motivated by the need for reliable confidence limits for multiple linear regression (MLR) parameters calculated from the shuttle data [Lukas and Firing, 1984, 1985] using McPhaden's [1982] model.

McPhaden [1982] used a seasonal MLR model consisting of a mean, linear trend, and annual and semiannual harmonics to analyze a time series sampled irregularly at 1- to 4-week intervals. He showed that occasional gaps in weekly wind observations substantially affected the estimated trend and seasonal cycle when important peaks of short duration were missed.

All such errors due to inadequate sampling are now commonly referred to as aliasing, although the term arose in the narrower context of Fourier analysis of evenly spaced data. In this original sense, sinusoids are aliases of each other if they are indistinguishable given a particular sample spacing. Any sinusoid with frequency greater than the Nyquist frequency (one half cycle per sample interval) will be seen as its alias in the frequency interval from zero to Nyquist. The seriousness of this type of aliasing depends critically on the periodicity of both the process and the sampling.

Shapiro and Silverman [1960] studied aliasing with different schemes of nonperiodic sampling. Leakage was avoided by considering infinite record lengths and therefore an infinity of samples. They showed that aliasing persists when periodic sample times are perturbed by random variations ("jitter").

The sample times are independent random variables, but the sample intervals are not. Aliasing vanishes (the power spectrum is uniquely determined) if sampling is completely aperiodic, and each sample interval is an independent random variable.

Aliasing effects and sampling requirements can be quantified by subsampling continuous time series. Luther and Harrisson [1984] subsampled island wind data to study errors in spectra calculated from monthly means of occasional ship reports. They found that the mean subsampled spectrum was the true spectrum plus white noise. For most of the wind records considered, three observations per month were adequate for the annual period, but at least 32 observations per month were needed to resolve the month-to-month differences after removal of the annual harmonic.

In this paper we analyze the sampling in the shuttle experiment. The shuttle data sets are described in section 2. Leakage and aliasing in the MLR model are considered in section 3. Monte Carlo experiments are used in section 4 to determine error bars for the seasonal model parameters and to explore sampling requirements. The experimental results are related to MLR theory in section 5, and conclusions are summarized in section 6.

2. DATA

The data used in this study were obtained during the NORPAX Hawaii-to-Tahiti Shuttle Experiment [Wyrtki et al., 1981]. The observations included shipboard hydrographic and current profile measurements, moored current meter and wind records, and island sea level data.

2.1. Shipboard Measurements

There were 15 cruises during the period February 1979 to June 1980. On all but the first, meridional sections crossed the equator near 150°W, 153°W, and 158°W. Conductivity, temperature, and depth (CTD) measurements were made from the surface to 1000 m at 1° latitude intervals [Wyrtki and Kilonsky, 1984]. XBT profiles to 450 m were made midway between CTD stations. Relative current profiles to 500 m were
Fig. 1. The response of the seasonal model to a sinusoid of unit amplitude is shown as a function of frequency from 0 to 6 cycles/yr. The heavy (light) curve is for the model with trend included (omitted). Record length varies from 11 to 24 months. As the amplitudes of the annual and semiannual harmonics depend on the phase of the input, both maximum and minimum amplitudes are shown.

measured at CTD stations from 10°N to 4°S and halfway between CTD stations within 3° of the equator [Firing et al., 1981]. Routine weather observations were recorded at each hydrographic station.

2.2. Moorings

A moored current meter (MCM) array was maintained near the equator, 153°W, from April 24, 1979, to June 2, 1980 [Knox and Halpern, 1982]. Hourly means of temperature, water velocity, and wind stress (calculated from buoy wind velocity measurements) were kindly provided by David Halpern. In this study we use records of temperature and velocity from 15-m depth at 0°40′N, 100 m at 0°40′S, and 250 m on the equator. We use the wind record from the buoy at 0°40′S. All of these records are continuous with the exception of two 1-week gaps during which the moorings were being replaced and an additional loss of velocity data from the 100-m record for two weeks at the end of the second deployment.

2.3. Island Measurements

Low-pass filtered daily mean sea level measurements from tide gauges at Fanning (4°N, 159°20′W) and Christmas (2°N, 157°30′W) islands were generously provided by Klaus Wyrtki. Daily mean subsurface pressure-temperature records from 5- to 10-m depth at Jarvis (0°23′S, 160°W) and Malden (4°S, 155°W) islands were also used. These records are continuous for the entire shuttle experiment.

Lacking continuous records of daily dynamic height, we
consider detided sea level as the best available surrogate. Barotropic tides dominate sea level variations, but dynamic height is influenced only by internal tides. As we do not know the relative contribution of the internal tides to sea level, we simply remove all high-frequency tides. Thus our calculations of dynamic height aliasing will be biased low.

Comparisons between sea level at an island and dynamic height measured nearby have shown a useful degree of similarity. For periods longer than 8 days, Wunsch [1972] found coherence of 0.4 to 0.5 between dynamic height relative to 1500 m near Bermuda and the residual sea level at Bermuda after removing tides and fluctuations correlated with wind and atmospheric pressure. He attributed some of the difference to barotropic fluctuations and some to local sea level perturbations due to the island topography. Wyrski [1980] compared Fanning Island sea level to dynamic height relative to 300 m, calculated from XBT data taken near the island and using a mean temperature-salinity relationship. He found 3.2 cm rms deviation from the regression between the two, similar to the typical variation found in either type of data over a period of a few days. On seasonal and longer time scales the two records appeared well correlated. The data were inadequate for calculation of coherence, but inspection of the plots suggested to Wyrski [1980] that the minimum period with useful coherence might be between a week and a month. Meyers [1982] found good agreement between island sea level and dynamic height in the western tropical Pacific on seasonal and interannual time scales. In summary, the relationship between island sea level and nearby dynamic height is still not well established, especially at periods of less than a month; but for longer periods, coherence is substantial.

3. The Seasonal MLR Model

We focus here on the seasonal MLR model used by McPhaden [1982].

\[ y(t) = B_0 + B_1(t - t_0) + B_2 \cos 2\pi \omega_d t + B_3 \sin 2\pi \omega_d t + B_4 \cos 4\pi \omega_d t + B_5 \sin 4\pi \omega_d t + \text{residual} \tag{1} \]

With \( \omega_d = 1 \text{ cycle/year} \) this model uses the annual and semiannual harmonics to approximate the seasonal cycle and a linear trend to model interannual variations. For all calculations in this paper the trend is centered; \( t_0 \) is the middle of the record, and \( B_0 \) is thus an estimate of the mean.

Aliasing and leakage of energy from neighboring frequencies due to finite record length, can be studied by calculating the response of a given MLR parameter to sinusoidal data of specified frequency and phase. If data are sampled from the function \( \cos(o t - \phi) \), then a scalar MLR parameter \( M \) will be

\[ M(\phi) = M_e \cos(\phi - \Phi) \]

where \( M_e \) and \( M_s \) are the responses to \( \cos \) and \( \sin \) respectively. The harmonic amplitudes are nonlinear functions of the MLR parameters. They are periodic functions of \( \phi / 2 \) with extrema

\[ A = \{1/2(CC^* + SS^*)[1 \pm (1 - Q)^{1/2}]\}^{1/2} \]

\[ Q = [2 \ln |(CS^*)/(CC^* + SS^*)|^2 \]

Here \( C \) and \( S \) are the complex amplitudes of a harmonic (e.g., \( B_2 + iB_3 \) for the annual harmonic) in response to \( \cos \) and \( \sin \), respectively, and the asterisk denotes the complex conjugate.

3.1. Leakage

To explore the leakage problem as a function of record length, we have calculated MLR parameter responses as functions of forcing frequency for monthly data with record lengths 11, 12, 13, 15, 18, and 24 months (Figure 1). The forcing frequency was incremented from 0 to \( \omega_n = 6 \text{ cycles/yr} \) in 0.1 cycle/yr steps. Since the trend is generally orthogonal only to the mean, leakage into other parameters depends strongly on whether the trend is included in the model. With short records, inclusion of the trend has a striking effect on the annual and semiannual harmonics. As expected, it reduces leakage to these harmonics from frequencies below annual. For the 11-, 12-, and 13-month records, however, this reduction of low-frequency leakage is at the expense of greatly increased leakage from higher frequencies, especially near 3 cycles/yr. With record lengths of 15 months or longer, the MLR including the trend becomes reasonably well behaved. The trend responds mainly to interannual frequencies and causes only slight leakage from higher frequencies into the annual harmonic. Note, however, that inclusion of the trend increases the response to the biannual harmonic when the record is 2 years.

Most shipboard time series from the Hawaii-to-Tahiti Shuttle Experiment are 15 months or longer, so the trend may be safely included in the MLR analysis. The moored current meter time series are only 13.5 months long; so leakage may be substantial in a seasonal model with trend included. It is not clear whether the benefit of the trend (reduced leakage from interannual frequencies) outweighs the disadvantage (increased leakage from higher frequencies).

3.2. Aliasing

Since aliasing is due to discrete, periodic sampling, we would expect it to be reduced by irregular sampling [Thompson, 1971] such as in the shuttle experiment (Figure 2). To show this, we have calculated the seasonal model frequency response for a variety of sampling schemes (Figure 3). Four of the samplings are based on the equator station times during the Hawaii-to-Tahiti Shuttle. The 150ºW, 153ºW, and 158ºW samplings used equator crossings at the respective longitudes, while the "all longitudes" sampling includes all equator crossings, irrespective of longitude. The sampling patterns labeled "even" are periodic. For example, the "all longitudes—even" pattern has the same record length and number of samples as "all longitudes," but the interval between samples is constant. The daily mean analysis is given first as a reference. It shows that increased sampling density reduces aliasing by raising the

---

**Fig. 2.** Sampling during the NORPAX Hawaii-to-Tahiti Shuttle. Times are in days from January 1, 1979. Vertical lines show equator crossing times, and dots indicate the subset of these times during which moored current meters were deployed. This subset was used for the Monte Carlo experiments (section 4).
Fig. 3. The response of the seasonal model to a sinusoid of unit amplitude is shown as a function of frequency from 0 to 50 cycles/yr. The heavy (light) curves in the lower panels are the maximum (minimum) amplitudes of the annual and semiannual harmonics. Sampling scheme (explained in the text) and record length are indicated to the right of each curve.
Nyquist frequency, but leakage remains because of limited record length.

With even sampling, perfect aliases of the mean and the annual and semiannual harmonics occur near integral multiples of the sampling frequency, starting at 11 cycles/yr for one longitude and 36 cycles/yr for combined longitudes. With irregular sampling the response is different. At 153°W and 158°W the first alias is at the same frequency as with even sampling, but the response is slightly reduced. At 150°W the first alias is near 6 cycles/yr, and the response to frequencies near 11 cycles/yr is conspicuously low. This is because 150°W samples are grouped in bimonthly pairs. At higher frequencies the response fluctuations appear random, with the possible exception of the unexpected peaks near 16 cycles/yr for 158°W. The maximum response above Nyquist frequency can approach unity, but the average levels are much lower.

The sampling with all longitudes together is considerably more effective in suppressing aliases. There are no aliases with unit amplitude, and there is no response peak near 36 cycles/yr for 11 cycles/yr. It is caused by the monthly clustering of samples evident in Figure 2. At higher frequencies the average response is lower for all longitudes together than for a single longitude. This is just a consequence of the familiar statistical result that the standard error of a sample mean varies inversely with the square root of the number of samples.

The reduction in response at alias frequencies due to irregular sampling is achieved at the expense of increased response at other frequencies. Periodic sampling preserves the regularity of an unresolved periodic signal, mapping it onto a single frequency below Nyquist. Irregular sampling randomizes the signal, scattering its variance into a variety of other frequencies.

4. MONTE CARLO EXPERIMENTS

The MLR method of curve fitting includes convenient estimates of uncertainty in the parameters (see the appendix). These estimates are based on assumptions about the model and the data which may not be valid. An alternative way of estimating parameter uncertainty is to perform Monte Carlo experiments in which either the sampling or the data set is varied randomly while parameter statistics are collected. This procedure requires a statistical model of either the data set or the sampling scheme. We choose the latter approach, as we are interested here in how well the shuttle sampling allows us to determine model parameters during the shuttle time interval, not in how well those parameters characterize conditions during a longer interval.

4.1. A Statistical Sampling Model

Any time series generated by sampling a continuous function of time during a particular interval can be considered as a member of an ensemble of possible outcomes of a random process. The sampling process of principal interest here will be referred to as "jittery"; consecutive sample times are independent random variables within successive evenly spaced intervals of uniform length. (This is the "jittered periodic" scheme of Shapiro and Silverman [1960].) The sampling method can
be stated mathematically as follows: If the domain $t_b \leq t \leq t_e$ is to be sampled $n$ times, choose the samples $t_i$ from $n$ segments of length $L$ centered at times $t_i$:

$$
t_i = t_0 + i \Delta t
$$

Each $g_i$ is an independent random variable with uniform probability density in the interval $-L/2 < g_i < L/2$ and zero probability density elsewhere. (This distribution is chosen more for mathematical convenience than for realism.) If $L = \Delta t$, then the entire domain may be sampled; otherwise, there will be portions of it between segments and at the ends that will be neglected.

Jittery sampling models manual data collection, such as the hydrographic and current profile measurements of the NORPAX shuttle. (Shipboard meteorological observations are typically made at fixed times, so sampling is periodic rather than jittery.) Although an attempt may be made to plan a series of identical cruises of equal length with evenly spaced starting times, vagaries of weather, logistics, and scientific interest will inevitably disturb the schedule.

4.2. Experiments

To determine realistic confidence limits for MLR parameters calculated from the shuttle data sets, we have performed a set of Monte Carlo experiments with the jittery sampling scheme. For each data set the record length was held constant at 406 days (13.5 months, the length of the MCM record) while the number of samples varied, taking the arbitrary values 8, 12, 18, 27, 40, and 60. All results given here are for $L = \Delta t$; the entire record was sampled. Gaps in the data were filled with the nearest data point. Each ensemble contained 100 realizations of the subsampling process. Experiments with up to 800 realizations showed that a sample of 100 is more than adequate for stable statistics.

We define the "correct" value of a parameter to be the limit as the number of evenly distributed samples becomes infinite within the time interval of interest. In practice, we use the parameter calculated from daily mean data as our standard. The difference between this standard and the Monte Carlo mean for a given sampling process is the bias.

For each parameter the rms bias calculated from all data sets and sampling schemes was 15–20% of the standard deviation. Hence the standard deviation of the parameter is the main source of uncertainty in any given MLR calculation; a 20% bias increases the rms error by only 2%.

4.3. Results

Results from the Monte Carlo experiments are given in Figures 4–8. Three factors should be kept in mind while interpreting these figures. First, for each experiment the standard deviations of parameters other than the mean depend on the distribution of variability among segments of the record [Draper and Smith, 1981]. For example, the trend is more sensitive to the ends of the record than to the middle. Second, the phase standard deviations are inversely proportional to the amplitudes (see the appendix). That is, they measure the noise to signal ratio for the harmonics. Third, the 406-day records used here are significantly shorter than the shuttle shipboard time series (Figure 2), so leakage is greater with the former (as shown in section 3). Hence the standard deviations in Figures 4–8 probably overestimate the errors in seasonal models of the shipboard time series.

Subject to these limitations, we will use Figures 4–8 to reach some specific conclusions about sampling requirements for adequate determination of the seasonal model parameters. We will consider a 10° standard deviation for phase as the maximum for "adequate" determination of a harmonic. This corresponds to 95% confidence limits of approximately ±20° and to a 17% relative standard deviation for amplitude (if the sine and cosine parameters are uncorrelated and have equal variance). Our conclusions are as follows:

1. Wind (Figure 4) must be sampled more than once per week to adequately determine the annual cycle (using 10° standard deviation of phase as the criterion). The mean zonal wind speed is determined to ±10% by a sample every 3 weeks, but more than weekly sampling is needed to measure the meridional component to the same relative accuracy. This is not surprising, since the mean meridional component is an order of magnitude smaller than the zonal component. The trends of both components are small in relation to shorter-period variations and cannot be distinguished from zero even with weekly sampling.

2. Daily mean sea level (Figure 5) must be sampled bi-weekly to determine the annual cycle. An annual mean and a trend can be determined to ±1 cm and ±2.5 cm/yr, respectively, with monthly sampling.

3. Sampling required to determine the annual cycle in temperature (Figure 6) varies strongly with depth, as the seasonal signal is most prominent near the surface. A sample every 3 weeks is adequate in the mixed layer, but weekly
The recommended sampling rates for sea level are of little value in themselves, as sea level is almost always sampled many times per day. They do, however, provide probable lower bounds on the sampling rates required for dynamic height of the sea surface relative to a deep reference level, as calculated from single CTD or XBT profiles. Faster rates might be required because of noise caused by internal tides. Dynamic height at depths other than the surface also might require different sampling rates, which we cannot determine from sea level data.

The general conclusion that we now draw regarding the Hawaii-to-Tahiti Shuttle Experiment is that the attempt to resolve zonal variability failed. On each of the three longitudes there were 11 or 12 samples during the 406-day interval used for the Monte Carlo experiments, for an average sampling interval of just over a month. Combining all three longitudes puts 35 samples in the same interval, for a 12-day average sampling interval. Monthly sampling was rarely adequate for defining a mean, much less the trend or annual cycle, whereas weekly to biweekly sampling was sufficient for many, though not all, of the variables.

5. **DISCUSSION**

Comparison of the experimentally determined parameter standard deviations to the MLR standard error estimates (Figures 4–8) shows the latter to be biased high, in some cases by more than a factor of 2. The bias tends to increase with the number of samples. Now we will show how this can be understood as a result of lack of fit in the MLR model.

To apply MLR theory to the Monte Carlo experiments in section 4, we assume that the functions used to model the data are slowly varying; we may consider them constant in any interval of length $L$. Then we may replace $t_i$ with $r_i$ in (A2), so each subsampling in a single Monte Carlo experiment is a
realization of a single random process; only the residuals change from one realization to the next. Now the residual vector \( \mathbf{R} \) can be split into an ensemble mean \( \overline{\mathbf{R}} \) and a deviation \( \mathbf{R}' \):

\[
\mathbf{R} = \overline{\mathbf{R}} + \mathbf{R}'(t)
\]

\[
\overline{\mathbf{R}}_n = \frac{1}{L} \int_{n-1/2}^{n+1/2} r(t) \, dt
\]

Jittery sampling ensures that the elements of \( \mathbf{R}' \) are uncorrelated, but the variance may differ substantially from one element to the next. Hence the common simplifying assumption, \( \mathbf{V}(\mathbf{R}) = S^2 I \), is not perfectly satisfied by our statistical model; weighted least squares would provide a better fit than (A2). Otherwise, the model is now identical to those treated in standard MLR theory [e.g., Draper and Smith, 1981]. The segment mean \( \overline{\mathbf{R}} \) is the systematic component of the residual, due to lack of fit of the model, and the segment deviation \( \mathbf{R}' \) is the random component.

Lack of fit has important consequences: (1) comparing one realization of a subsampling process with another realization, the one with the smaller estimated parameter errors (from (A3) and (A4)) is likely to have the larger actual parameter errors; (2) the internal error estimates are biased high. For a given time series the bias increases with the sampling frequency, as more of the residual variance goes into \( \overline{\mathbf{R}} \) and less into \( \mathbf{R}' \). High frequencies will contribute variance primarily to the segment deviations, but variance from lower frequencies will go to the segment means (see the appendix). This situation accounts for the large differences in bias among records. For example, Jarvis sea level estimated errors (Figure 5) are biased much higher than those for wind stress (Figure 4). Figure 9 shows that most sea level residual variance is at frequencies below the sampling rate, whereas much wind residual variance is at higher frequencies.

**6. SUMMARY**

The NORPAX Tahiti Shuttle Experiment was designed to study seasonal time scales in currents and hydrography during a 15-month period. We have concluded that the shipboard sampling was adequate for this purpose only if the longitudinal separation of the three meridional sections is ignored and only for temperature and the zonal velocity component. Much more frequent sampling would be required to determine a significant mean or seasonal cycle in the meridional velocity component.

Analysis of the seasonal MLR model and Monte Carlo subsampling experiments has led to some more general conclusions as well.

1. Inclusion of the trend in a seasonal model is of dubious value if the record length is 13 months or less. This confirms the commonsense notion that more than one realization is needed to determine a seasonal cycle amid interannual and other fluctuations.

2. The irregularity of the shipboard sampling substantially reduces aliasing of single high-frequency harmonics into the seasonal model. This effect is important in practice only if the spectrum contains significant peaks.

3. Parameter error estimates generated internally by the MLR technique are biased high by lack of fit in the model.
Fig. 8. Standard deviations of seasonal model parameters for meridional velocity component at 15-, 100-, and 250-m depth; see Figure 4 caption.

Hence they tend to give an overly pessimistic view of the significance of the parameters. This tendency increases with the proportion of variance in the record that is below the sampling frequency.

**APPENDIX**

**Review of the MLR Method**

Equation (1) may be written more generally as

\[
Y = XB + R
\]

where \( Y \) contains \( n \) samples taken at times \( t_i \), \( F(t) \) is a vector of \( p \) functions of time that model \( Y \), and \( R \) is a vector of residuals sampled from a function \( r(t) \). Then \( b \), the value of \( B \) that minimizes \( R'R \) (least squares), is

\[
b = (X^TX)^{-1}X^TY \]

(\( X^T \) denotes the transpose of a matrix, and the superscript \(-1\) denotes the inverse. The element \( Z_{ij} \) is the sensitivity of parameter \( j \) to data point \( i \). If the MLR model fits the process that generated the data and if \( E(RR^T) = \sigma^2 I \), where \( E \) denotes expected value over an ensemble of realizations and \( I \) is a unit matrix, then \( b \) is a random variable with variance-covariance matrix

\[
V(b) = (X^TX)^{-1}\sigma^2
\]

(\( A3 \))

An unbiased estimator of \( \sigma^2 \) in this simple case is the residual variance,

\[
s^2 = \frac{(Y - Xb)^T(Y - Xb)}{(n - p)} \quad (A4)
\]

Expressions (A3) and (A4) provide the standard internal error estimates for the MLR parameters.

**Standard Deviations of Amplitudes and Phases**

Amplitude and phase are nonlinear functions of the sine and cosine coefficients. Given normal distributions of sine and cosine coefficients, the distributions of amplitude and phase are non-Gaussian, and exact calculations of variance are not straightforward. To avoid this difficulty, we use a linearized approximation [Jenkins and Watts, 1968]:

\[
\begin{align*}
\text{Var}(A) &= C^2 \text{Var}(C) + S^2 \text{Var}(S) + 2CS \text{Cov}(C, S) \\
\text{Var}(\phi) &= (C^2 - \sigma^2 + S^2)^2 (A6)
\end{align*}
\]

where \( C \) and \( S \) are mean sine and cosine coefficients, \( A \) and \( \phi \) are the amplitude and phase, and \( \text{Var} \) and \( \text{Cov} \) denote variance and covariance, respectively. For internal MLR error estimates the variances and covariances on the right-hand side of (A5) and (A6) are estimated from (A3). In the Monte Carlo experiments they are calculated from each ensemble of sine and cosine coefficients.

This linearization is valid when the range of variation of \( C \) and \( S \) is small in comparison with the mean amplitude. When
This condition is not met, \((A6)\) can give absurdly large values. Then there is no usable phase information; so the inapplicability of the linearization is of little practical consequence.

### Spectral Partition of Residual Variance

The residual function \(r(t)\) can be described over the domain \(0 \leq t < T = nL\) by its Fourier series:

\[
r(t) = \sum_{j=0}^{\infty} a_j \exp(ia_jt)
\]

where \(a_j = 2\pi j / T\), \(a_j\) is complex, and it is understood that the real part of the right-hand side is to be taken. The variance of \(r(t)\) is

\[
\text{Var}[r(t)] = 1/T \int_0^T [r(t) - a_0][r(t) - a_0]^* dt
\]

\[
\text{Var}[r(t)] = \sum_{j=1}^{\infty} a_j a_j^*
\]

Applying the definition \((5)\) to \((7)\) gives

\[
\tilde{r}_i = \sum_{j=0}^{\infty} a_j \exp(i\omega_j t) \frac{\sin(\omega_j L/2)}{\omega_j L/2}
\]

If the segment mean is defined as a discontinuous function of time, with the constant value \(\bar{r}_i\) in the \(i\)th segment, then \((8)\)

\[
\text{Var}[r(t)] = \sum_{j=1}^{\infty} a_j a_j^* \left( \frac{\sin(\omega_j L/2)}{\omega_j L/2} \right)^2
\]

\[
\text{Var}[r(t)] = \sum_{j=1}^{\infty} a_j a_j^* \left( \frac{\sin(\omega_j L/2)}{\omega_j L/2} \right)^2
\]

This specifies the total variance in the segment deviations but not the distribution among the different segments. The latter involves all possible cross products of the \(a_j\), since the complex exponentials in \((7)\) are not orthogonal over a single segment.

The first zero crossing of the \((\sin x)/x\) factor in \((10)\) and \((11)\) is at \(\omega = 2\pi / L\), which is twice the Nyquist frequency. At the Nyquist frequency, \((\sin x)/x = 0.637\); so significant energy goes into both the segment mean and the deviation. At about 1.6 times the Nyquist frequency, \((\sin x)/x = 0.21\), a value that is not exceeded at any higher frequency. It may be taken as the highest frequency at which significant variance goes into the segment means.

### Acknowledgments
We are grateful to D. Halpern and K. Wyrtki for providing essential data. We thank S. Murakami for her patience and proficiency in programing and graphics. Helpful comments by J. Picaut, M. McPhaden, and S. Hayes on an earlier version of the manuscript are appreciated. Funding for this work was provided by the National Science Foundation under grant OCE82-08944. Contribution numbers 1647 of the Hawaii Institute of Geophysics and JIMAR 84-0089 of the Joint Institute for Marine and Atmospheric Research.

### References


Wyrtki, K., Sea level during the NORPAX Test Shuttle Experiment, Science, 211, 22–28, 1981.
