## ON THE STATISTICS OF AMBIENT NOISE

### Authors
Elizabeth M. Arase and T. Arase

### Performing Organization
Columbia University
Hudson Laboratories
3055 Broadway, New York, New York 10027

### Contract or Grant Number
Nonr-266(84)

### Distribution Statement
Approved for public release; distribution unlimited.

### Security Classification
UNCLASSIFIED
ON THE STATISTICS OF AMBIENT NOISE

by

Elizabeth M. Arase
and
T. Arase

Hudson Laboratories of Columbia University
Dobbs Ferry, New York 10522

* Hudson Laboratories of Columbia University Informal Documentation
No. 126.

† This work was supported by the Office of Naval Research.
On the Statistics of Ambient Noise, Elizabeth M. Arase and T. Arase. Hudson Laboratories of Columbia University, Dobbs Ferry, New York 10522. The statistics of ambient noise have been investigated previously for single receivers (A. H. Green, Bell Tel. Lab. Tech. Report No. 10, 1962). No studies have been made for ambient noise with arrays. Data are presented for ambient noise measured with arrays with 30 to 60 elements. Amplitude samples of ambient noise were taken at 30-msec intervals. The period is long enough to ensure that successive samples are independent. Sets of 2000 to 3000 points were taken. Noise statistics were also taken with random addition of the elements. No significant difference was found for the statistics of the two cases. Distribution functions and moments up to the fourth order were computed and will be presented. (Hudson Laboratories of Columbia University Informal Documentation No. 126. This work was supported by the U. S. Office of Naval Research.)
The statistics of ambient noise have previously been investigated for single receivers in the frequency range to about 120 Hz. No studies have been made for ambient noise with arrays, or for single receivers at a higher frequency. We have studied the statistics of ambient noise with a large number of receivers (from 30 to about 60), which formed an over-spaced array in the frequency range 300 to 500 Hz, and was steered for RSR arrival.

The large number of receivers were phased by measuring the arrival time of the RSR downgoing arrival from a distant pulsed source. The pulses were short enough so that a particular arrival could be identified and strong enough so that the signal was much larger than the noise.

The receiver signals were clipped, delayed, and added to form beams. The correctness of the beams formed was partly determined by measuring the signal-to-noise ratio at the output of the beamformer for short pulsed signals. For high signal-to-noise cases, the signal power adds as $N^2$ where $N$ is the number of elements, and the noise power adds as $N$. Consequently the output signal-to-noise should be proportional to $N$, as may be seen from Fig. 1 to be the case for the results presented here.

After a best beam had been obtained on pulsed signals, the source was turned off and these same delays were used on ambient noise. Some data were taken on tape for off-line processing in order to compare random delays with a steered beam for the same sample of data. Only such records were used, in which no nearfield sources were apparent, such as whale noise, overhead ships, or rainstorms.

The amplitude output was recorded on a Visicorder, and a typical portion of such a record is shown in Fig. 2. Samples of the first two sets with steered and random delay, called hereafter sets I, II, III, and IV, were
read at time intervals of 30 msec. A total of 2000 points was read for each record, corresponding to a sample length of 1 min. For the last set, called set X, dependent samples were taken at 1 and 3 msec after the first, to determine the consistency of dependent sets. For each of these dependent sets we took about 3000 sample points, corresponding to a record length of 1.5 min.

Since the system is clipped, the distribution would be expected to be binomial, if the unclipped distribution is Gaussian. However, for the number of elements used, between 37 and 64, the difference between the two distributions is insignificant.

Figures 3 and 4 show the cumulative distribution as a function of the deflection in cm as well as in volts for random delays and for a steered array. The curves are plotted on probability paper, on which a normal distribution appears as a straight line. Figure 5 shows the three dependent sets, which seem to vary as much from each other as the steered and random sets do in the previous figures. Certainly, grossly, all of the distributions have the features of a normal distribution. However, better tests have to be undertaken to determine their normality. One method consists of computation of the moments. Figure 6 shows the nth moment for r sample points, where $\bar{x}$ is the mean of the sample. The coefficient of skewness $y_2$, which is sometimes defined as $y_2^2$, as for example by Pearson, is zero for a normal distribution. The peakedness or kurtosis is equal to 3. Figure 7 shows the computed values of $y_2$ and $\beta_2$ for the experimental distributions. Pearson has computed the confidence limits for $y_2$ and $\beta_2$ as function of the number of data points. The 99 percent confidence interval for 2000 points is given at the bottom of the figure, and we see that set IV and sets X and Y fall outside of these limits, as indicated by the •.
The moments tests have the disadvantage that large amplitude values are heavily weighted. Quite often experimental distributions which are normal in the central region have tails, and are rejected for this reason, not always validly. One test for normality, which takes care of this objection, is the Kolmogorov-Smirnov test, which can also be used to test for the stationarity of our distribution. Figure 8 shows the equations for this test.

We find the maximum difference between the theoretical distribution $F(x)$, based on the experimental value of the average value of $x$ and the standard deviation, and the experimental cumulative distribution $S_n(x)$. For this to lie within the 99 percent confidence interval, the maximum difference has to be less than a constant $a$, which is dependent on the number of data points. To test for stationarity between sets, the difference between the sets in each interval has to be less than another constant. We first test for normality of each set. Figure 9 shows the experimental step function, the central line which is the theoretical distribution with which the experimental distribution is compared, and the 95 and 99 percent confidence limits.

The experimental distribution is always less than the confidence limits. For the next few sets we shall only show the inset area enlarged, although we also tested in the outer regions. Figures 10, 11, and 12 show sets II, III, and IV. Of these sets, III and IV exceed the confidence interval. Set III had previously been shown to be normal from the moments test. The equivalent tests for sets X, Y, Z show that all of these sets exceed the 99 percent confidence interval.

To test for stationarity of our sets, we divided each set into four parts, obtaining subsets 1, 2, 3, and 4 with 500 points each for sets I, II, III, and IV. The individual sets again were tested for normality, and then for stationarity according to the Kolmogorov-Smirnov test. The maximum inter-set
coefficients are given, and exceed the 99 percent limits in all but in set IV. So in general, the distributions are nonstationary. This is more apparent then in Green's work, since he worked at lower frequencies, where ambient noise appears to be more stationary.

The last two figures (17 and 18) give a summary of the normality of the small as well as the large sets. The $R$ indicates the small as well as the large sets which were rejected in the moments test or in the Kolmogorov-Smirnov test. The majority of the small sets fall within the 99 percent confidence interval of normality. That a large number of total sets is rejected is probably due to the nonstationarity of the noise.

Further work for a single hydrophone in this frequency range is necessary to investigate the character of ambient noise.
REFERENCES


Fig. 1. Signal-to-noise ratio as function of number of elements.

Fig. 2. Typical Visicorder record.

Fig. 3. Sets I and II.

Fig. 4. Sets III and IV.

\[
\mu_n = \left( \frac{1}{r} \right) \sum_{i=1}^{r} (x_i - \bar{x})^n
\]

\[
y_2 = \frac{\mu_3}{(\mu_2)^{3/2}} = 0
\]

\[
\beta_2 = \frac{\mu_4}{(\mu_2)^2} = 3
\]

Fig. 5. Dependent sets X, Y, Z.

Fig. 6. Definition of moment, coefficient of skewness, and peakedness.
AMBIENT NOISE STATISTICS

<table>
<thead>
<tr>
<th>#ELEMENTS</th>
<th>$y_2$</th>
<th>$\beta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEERED</td>
<td>-0.10</td>
<td>3.25</td>
</tr>
<tr>
<td>RANDOM</td>
<td>-0.12</td>
<td>2.93</td>
</tr>
<tr>
<td>STEERED</td>
<td>+0.01</td>
<td>3.20</td>
</tr>
<tr>
<td>RANDOM</td>
<td>-0.08</td>
<td>3.61</td>
</tr>
<tr>
<td>STEERED</td>
<td>-0.16</td>
<td>3.13</td>
</tr>
<tr>
<td>RANDOM</td>
<td>-0.12</td>
<td>2.88</td>
</tr>
</tbody>
</table>

99% $-0.127 \leq y_2 \leq 0.127$
2.77 $\leq \beta_2 \leq 3.28$

Fig. 7. Computed coefficients of skewness and peakedness.

Kolmogorov Smirnov Tests

$D_n = \sup_{-\infty < x < +\infty} \left| F_n(x) - S_n(x) \right|$

$< \frac{\lambda}{\sqrt{n}} = \alpha$

$D_n = \sup_{-\infty < x < +\infty} \left| S_{n_1}(x) - S_{n_2}(x) \right|$

$< \frac{\lambda \sqrt{n_1 n_2}}{n_1 + n_2}$

Fig. 8. Kolmogorov Smirnov tests.

Fig. 9. K. G. Smirnov test.

Fig. 10. K. G. Test, Set II.

Fig. 11. K. S. Test, Set III.

Fig. 12. K. S. Test, Set IV.
Fig. 13. Stationarity Test.

**SET I**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>560</td>
<td>300</td>
<td>970</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>420</td>
<td>931</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

99% $\alpha = 1031 \times 10^{-4}$
95% $\alpha = 660 \times 10^{-4}$

Fig. 14. Stationarity Test.

**SET II**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>630</td>
<td>1060</td>
<td>660</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>940</td>
<td>340</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

99% $\alpha = 1031 \times 10^{-4}$
95% $\alpha = 660 \times 10^{-4}$

Fig. 15. Stationarity Test.

**SET III**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>600</td>
<td>360</td>
<td>1100</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>340</td>
<td>720</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

99% $\alpha = 1031 \times 10^{-4}$
95% $\alpha = 660 \times 10^{-4}$

Fig. 16. Stationarity Test.

**SET IV**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>360</td>
<td>320</td>
<td>465</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>260</td>
<td>292</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

99% $\alpha = 1031 \times 10^{-4}$
95% $\alpha = 660 \times 10^{-4}$

Fig. 17.

**X**

<table>
<thead>
<tr>
<th></th>
<th>XOM</th>
<th>XOM</th>
<th>XOM</th>
<th>XOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>R</td>
<td></td>
<td>R</td>
</tr>
</tbody>
</table>

99% $\alpha$

Fig. 18.