Modeling the Stationarity and Gaussianity of Spontaneous Electroencephalographic Activity

JAMES A. McEWEN, STUDENT MEMBER, IEEE, AND GRANT B. ANDERSON

Abstract—**Considerable motivation exists for the development of an adequate statistical model for spontaneous electroencephalographic (EEG) activity. At present, almost all methods of time-domain and frequency-domain EEG analysis are based on implicit assumptions regarding the statistical characteristics of the underlying random process, particularly with respect to the extent of stationarity and Gaussianity of the process. However, the actual characteristics of specific EEG ensembles have not been extensively investigated. In this paper, a technique is proposed for estimating the degree of wide-sense stationarity and the degree of Gaussianity of an ensemble of EEG records. Results which have been obtained by applying this technique to three relatively large ensembles of multichannel EEG data are also described. In addition, the comparative advantages of employing alternate methods of EEG analysis are discussed in relation to the estimated degree of stationarity and Gaussianity of the particular EEG ensembles under consideration.**

INTRODUCTION

CONSIDERABLE motivation exists for the development of an adequate statistical model for spontaneous electroencephalographic (EEG) activity. At present, almost all methods of quantitative EEG analysis are based on certain implicit assumptions regarding the statistical characteristics of the underlying random process, particularly with respect to the extent of stationarity and Gaussianity of the process. The efficacy of alternate analytic techniques depends upon the degree to which such assumptions are justified by the characteristics of the particular ensemble of EEG segments being analyzed. In addition, a better understanding of some of the statistical properties of different EEG ensembles might eventually result in a better understanding of the neurophysiological mechanism of spontaneous EEG generation, a mechanism which is still not well understood.

Despite such motivation, relatively few investigations of the statistical properties of specific EEG ensembles have been described in the literature. The first studies of the EEG amplitude probability distribution suggested a striking similarity to the normal or Gaussian distribution [1], [2]. A later analysis of one 8.33 sec EEG segment from each of four subjects also showed that in two cases the amplitude distributions closely fitted a Gaussian distribution [3]. However, subsequent reports by others contained rather contradictory results. For example, tests of thirty 52.8 sec EEG segments for Gaussianity resulted in 29 rejections; the investigators concluded that the spontaneous EEG could not be modeled as a normal random process because not even its amplitude distribution was Gaussian [4]. Elul suggested that this study illustrated an extreme case where non-stationarity of the EEG was erroneously construed as indicative of a non-Gaussian distribution [5]. He tested successive 2 sec EEG segments from one subject and reported that the EEG was Gaussian 66 percent of the time in the resting state, shifting to 32 percent during a mental arithmetic task. Although the results of some later studies appear to agree with those of Elul (e.g., [6]), others do not. For example, Dumermuth *et al.* commented that most of the 40 sec EEG segments which they had analyzed deviated from Gaussianity [7]. Following the suggestion of Elul, they also analyzed 4 sec EEG segments in an attempt to reduce effects due to non-stationarity, but reported even stronger deviations from a Gaussian model [8], [9].

Several factors can be identified which have contributed to the previously described inconsistencies in the literature. Many early investigations involved relatively small ensembles of EEG segments from very few subjects. Frequently, EEG data from only one non-standardized channel were considered. The reliability and comparability of the results obtained in such studies were therefore affected by topological differences, by statistical variability due to small sample sizes and by inter-subject EEG variation. Another factor contributing to discrepancies among published findings concerns the different EEG digitization rates which were used: it will be shown in this paper that different sampling rates change the efficacy of statistical hypothesis tests. Finally, the problem of estimating the degree of stationarity of a particular ensemble of EEG segments has seldom been considered directly in such investigations. Attempts were instead made to circumvent the problem of stationarity when investigating Gaussianity by subdividing the EEG into very short segments in the expectation that any non-stationary effects would be reduced.

In this paper, a technique is proposed for estimating the degree of wide-sense stationarity and the degree of Gaussianity of an ensemble of EEG records. Results which have been obtained by applying this technique to three relatively large ensembles of multichannel EEG data are also described. In addition, the comparative advantages of employing alternate methods of EEG analysis are discussed in relation to the estimated degree of stationarity and Gaussianity of the particular EEG ensembles under consideration.

THEORETICAL CONSIDERATIONS

Random Process Characterization

The ensemble of all possible time functions which can be generated by a particular source together with their respective probabilities of occurrence defines a random process. Spontaneous EEG activity may therefore be modeled as a random

Manuscript received October 22, 1974.

The authors are with the Department of Electrical Engineering, University of British Columbia, Vancouver, B.C., Canada.

process. Any such process, denoted by $X(t)$, is said to be completely characterized or modeled if its nth order distribution function

$$F[x_1, \cdots, x_n;\ t_1, \cdots, t_n] = P[X(t_1) \leqslant x_1, \cdots, X(t_n) \leqslant x_n] \quad (1)$$

is known for any n and any set of sampling times $t_1, \cdots, t_n$ ([10], pp. 296-297). For most random processes it is difficult to obtain empirical estimates of (1). However, if a particular random process is both Gaussian and stationary then the problem of modeling it by estimating (1) is greatly simplified.

Briefly, a random process $X(t)$ is said to be Gaussian or normal if its nth order probability density function

$$f[x_1, \cdots, x_n;\quad t_1, \cdots, t_n]$$

obtained by differentiating (1) with respect to all variables x_i, takes the form of a jointly Gaussian distribution, i.e.,

$$f(x_1, \cdots, x_n;\ t_1, \cdots, t_n) = \frac{\exp\{-1/2([x] - [u])[K]^{-1}([x] - [u])^T\}}{(2\pi)^{n/2}(|K|)^{1/2}} \quad (2)$$

where

$$[x] = [x_1, \cdots, x_n]$$

$$[u] = [E\{X(t_1)\}, \cdots, E\{X(t_n)\}] = [u_1, \cdots, u_n]$$

$$[K] = \begin{bmatrix} k_{11} & \cdots & k_{1n} \\ \vdots & & \vdots \\ k_{n1} & \cdots & k_{nn} \end{bmatrix}$$

$$k_{ij} = E\{(x_i - u_i)(x_j - u_j)\} \quad (3)$$

and $|K|$ is the determinant of $[K]$, the covariance matrix ([11], pp. 111-112).

A random process $X(t)$ is said to be strictly stationary if none of its statistics are affected by a shift in time origin, i.e., if the two processes $X(t)$ and $X(t + \xi)$ have the same statistics for any ξ. A much weaker condition is that of wide-sense stationarity in a finite time interval: if

$$E[X(t)] = \mu = \text{constant}$$

and if the autocorrelation function is given by

$$R(t_i, t_j) = R(\tau), \qquad \tau = |t_j - t_i|$$

for all t, t_i and $t_j \in [0, T]$, then $X(t)$ is said to be wide-sense stationary in the interval $[0, T]$ ([10], pp. 300-304). Under this condition, (3) becomes

$$\begin{aligned} k_{ij} &= E\{(x_i - u_i)(x_j - u_j)\} \\ &= E\{x_i x_j\} - u_i u_j \\ &= R(t_i, t_j) - \mu^2 \\ &= R(\tau) - \mu^2 \end{aligned} \quad (4)$$

for all t_i, $t_j \in [0, T]$. From (2) and (4) it is therefore evident that, under the condition of wide-sense stationarity in the interval $[0, T]$, a Gaussian random process $X(t)$ is completely specified by its mean and autocorrelation function in the interval.

If it can be shown that a random process $X(t)$ is ergodic [12], then such statistics as the mean and autocorrelation function can be calculated from a single sample function, denoted by $x(t)$, i.e.,

$$E[X(t)] = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, dt = \mu \quad (5)$$

and

$$\begin{aligned} R(\tau) &= E[X(t)X(t + \tau)] \\ &= \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t + \tau)\, dt \\ &= R(\tau) \end{aligned} \quad (6)$$

where $R(\tau)$ represents the time autocorrelation function. However, an empirical test for ergodicity would require extensive ensemble calculations and would certainly not be feasible when only a limited number of sample functions of relatively short duration are available. Under these conditions ergodicity is usually assumed and any desired ensemble statistics are estimated from the individual characteristics of all available sample functions. For example, if all sample functions can be modeled as the output of a wide-sense stationary Gaussian process in an interval $[0, T]$, then the mean and autocorrelation function are sufficient statistical descriptors of the process in the interval. These ensemble descriptors can be estimated in practice by averaging the means and the time autocorrelation functions (or, equivalently, the power spectra) of the available sample functions.

A necessary requirement before any such modeling of observed EEG activity can be attempted is that some empirical procedures be established for testing individual EEG segments, at a specified significance level, for wide-sense stationarity and Gaussianity.

Tests of Wide-Sense Stationarity and Gaussianity

Assume that $[x_1, \cdots, x_{2n}]$ has been obtained by sampling a band-limited EEG signal $x(t)$ at or above the Nyquist rate during the time interval $[0, 2T]$. Although an exact determination of the degree of wide-sense stationarity and Gaussianity of $x(t)$ in the given interval is not possible, useful estimates of these statistical properties can be obtained by the application of certain hypothesis testing procedures.

A procedure for determining whether or not $[x_1, \cdots, x_{2n}]$ can be considered to be a set of samples from a wide-sense stationary function can be based on the requirement that the amplitude distributions and the power spectra calculated for the sample subsets $[x_1, \cdots, x_n]$ and $[x_{n+1}, \cdots, x_{2n}]$ must not be significantly different. Specifically, a test for the wide-sense stationarity of a given sample set can be constructed by first dividing the set into two equal subsets and calculating an amplitude histogram and power spectrum for each. Then the two-sample Kolmogorov-Smirnov (K-S) test [13], [14] can be employed to compare the sample amplitude and spectral distribution functions of each. The two-sample K-S test is based on the statistic D_2 which is defined as

$$D_2 = \sup_{\text{all } s} |F_n(s) - G_n(s)|$$

where $F_n(s)$ and $G_n(s)$ are distribution functions calculated from a set of samples of size n from populations F and G, respectively. A large value of D_2 resulting from application of the two-sample K-S test would indicate rejection, at some significance level, of the null hypothesis that F and G are identical. When $[x_1, \cdots, x_n]$ and $[x_{n+1}, \cdots, x_{2n}]$ are tested in this manner, rejection of either the hypothesis of identical amplitude distributions or the hypothesis of identical spectral distributions indicates that the original EEG signal cannot be modeled with confidence as a sample function of a random process that is wide-sense stationary over the interval $[0, 2T]$. Thus, rejection of either hypothesis for a given set of samples constitutes an empirical upper bound on the interval of wide-sense stationarity, i.e., in this instance the interval of wide-sense stationarity for the random process of which $x(t)$ is a sample function is assumed to be less than $2T$.

Testing the amplitude distribution of a set of EEG samples $[x_1, \cdots, x_{2n}]$ for Gaussianity or normality is accomplished by means of a goodness of fit test. The K-S goodness of fit test is employed because it has been shown that, with the population mean and variance estimated by the sample mean and variance, it yields a test for normality which is more powerful than the more popular chi-square test [13]-[15]. The K-S statistic D_1 represents the least upper bound of the differences between the empirical and assumed distribution functions:

$$D_1 = \sup_{\text{all } s} |F_{2n}(s) - F(s)|$$

where $F_{2n}(s)$ is the distribution function calculated from the set of $2n$ samples and $F(s)$ is the assumed distribution function. If D_1 is too large, the null hypothesis that $F(s)$ represents the population distribution function is rejected.

Experiment

In order to apply the previously described tests for Gaussianity and wide-sense stationarity to some actual ensembles of EEG activity, 30 hospital patients in the best surgical risk category were selected for EEG recording before and during general anesthesia. Two of the most commonly used anesthetic combinations were employed: 15 subjects underwent halothane/nitrous oxide anesthesia and 15 subjects received alphaprodine/nitrous oxide anesthesia. As illustrated in Fig. 1, EEG activity was recorded from two pairs of bilaterally symmetric, bipolar channels: F3-C3, F4-C4, C3-01 and C4-02, according to the International 10-20 System [16]. A Beckman 8-channel electroencephalograph, with its highpass filter set at 0.54 Hz and its lowpass filter set at 30.0 Hz to reduce artifact and interference, was connected to a Hewlett-Packard 3960 4-channel FM instrumentation tape recorder. Two 64 sec segments of EEG activity without visually obvious artifact were recorded from each subject while awake and resting with eyes closed, approximately one hour before surgery. Two additional EEG segments of 64 sec duration were later recorded from the same subject at a surgical level of general anesthesia. Three different sets of multichannel EEG segments were thus obtained for consideration: one set of 60 baseline

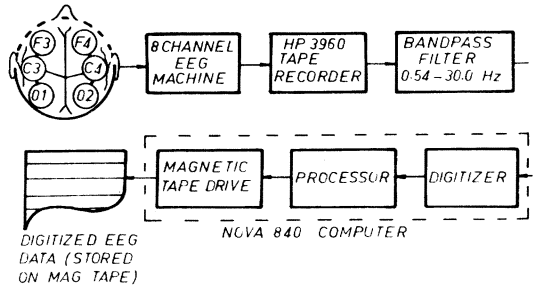

Fig. 1. General configuration of the EEG data acquisition system.

segments from 30 awake and resting subjects, a second set of 30 segments from 15 of these subjects during halothane/nitrous oxide anesthesia, and a third set of 30 segments from the other 15 subjects during alphaprodine/nitrous oxide anesthesia. Some typical EEG segments from each of these three sets of data are shown in Fig. 2. Further details regarding the data acquisition procedures and the anesthetic techniques are given elsewhere [17], [18].

After the acquisition of these three sets of EEG segments had been completed, it was desired to determine the best rate at which to sample and digitize the data. Because the EEG segments had already been lowpass filtered at 30.0 Hz, the theoretical minimum sampling rate, as given by the Sampling Theorem ([11], pp. 400-405), was 60.0 Hz, i.e., the Nyquist rate. The filter roll-off characteristics and the computational desirability of setting the sampling rate to a power of two indicated that the most practical minimum sampling rate, denoted by F_s, would be 64 Hz. Most of the previous investigations of Gaussianity or stationarity have considered EEG data sampled at rates of from $2F_s$ to $4F_s$ and even higher. However, statistical hypothesis tests such as the K-S and chi-square tests assume that the set of samples to be tested represents a set of statistically independent random variables or observations. Therefore, when this assumption of statistical independence is violated because of an unnecessarily high sampling rate, one can expect the efficacy of such tests to decrease accordingly.

To examine and illustrate the effect of different sampling rates on statistical hypothesis tests, 30 of the recorded 64 sec baseline EEG segments from C4-02 were reproduced, bandpass filtered from 0.54 Hz to 30.0 Hz with Krohn-Hite 3342R filters, and digitized at a rate of 512 Hz or $8F_s$. By considering every second, fourth or eighth sample, it was also possible to study EEG data with an effective sampling rate of $4F_s$, $2F_s$, or F_s. At each of these sampling rates a K-S goodness of fit test for Gaussianity, at the 0.05 significance level, was performed on each of the M available EEG segments of T sec duration, where

$$T = 2^i, \quad i = 0, 1, \cdots, 6 \tag{7}$$

and

$$M = \frac{30 \cdot 64}{T}. \tag{8}$$

The results of these tests are summarized in Fig. 3 and clearly indicate the desirability of using a sampling rate as little above the Nyquist rate as practicable.

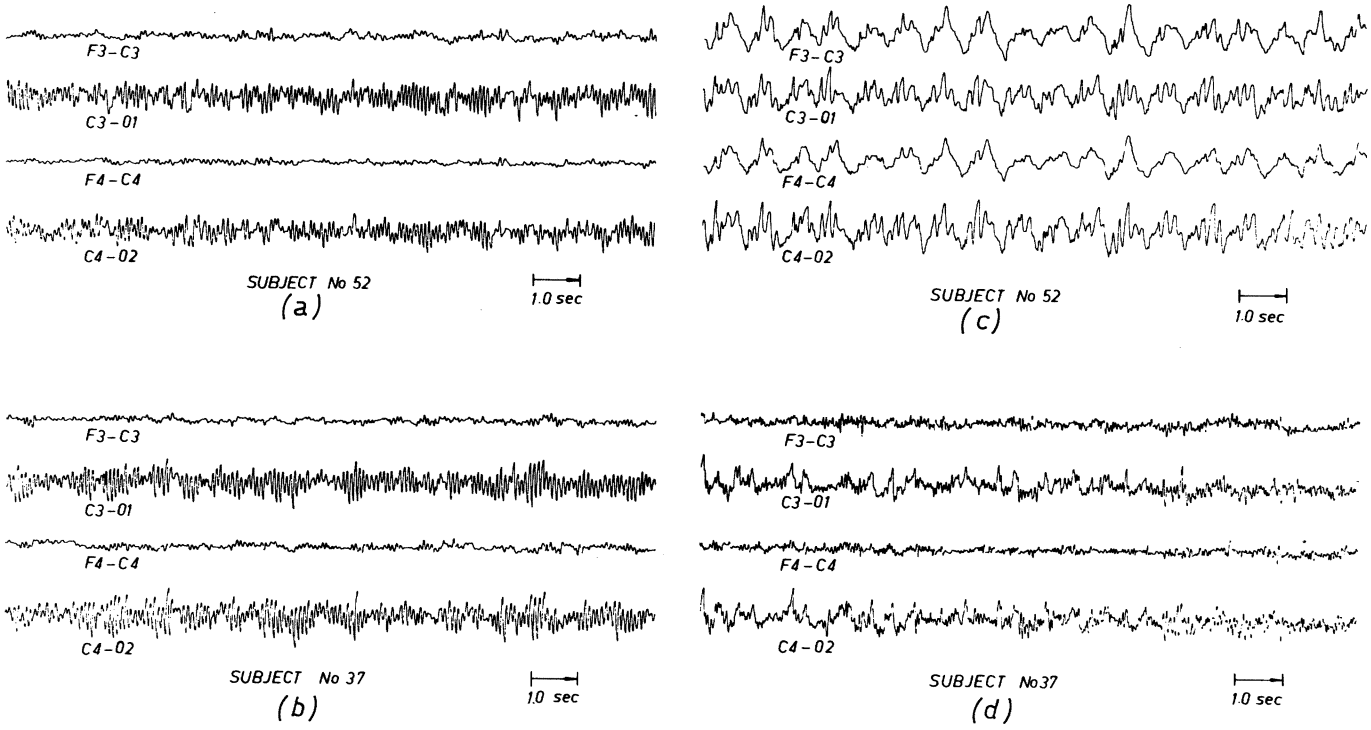

Fig. 2. Typical multichannel samples of spontaneous EEG activity from three different ensembles: (a) and (b) represent baseline EEG activity, (c) was recorded during halothane/nitrous oxide anesthesia and (d) was recorded during alphaprodine/nitrous oxide anesthesia.

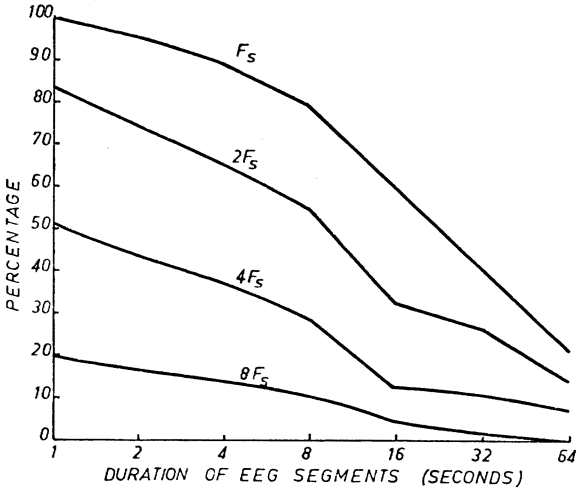

Fig. 3. Effect of increased sampling rates on K-S goodness of fit tests for Gaussianity. F_s is equal to 64 Hz, slightly above the Nyquist rate. The percentage of EEG segments of a specified duration which could be modeled as Gaussian is plotted for 4 different sampling rates.

To reduce error in the computation of power spectra, a sampling rate of 128 Hz was used to digitize all 120 EEG segments from the three ensembles under consideration. However, in view of the results in Fig. 3, EEG data with an effective sampling rate of 64 Hz were prepared by considering every second sample value, and were used to compute all sample amplitude distribution functions needed for the previously described tests for wide-sense stationarity and Gaussianity.

Recall that, for an EEG segment $x(t)$ to be modeled as a sample function of a process that is wide-sense stationary in the interval $[0, 2T]$, a necessary condition is that the amplitude distribution functions and the power spectral distribution functions of $x(t)$ in the intervals $[0, T]$ and $[T, 2T]$ must not be significantly different. The distribution functions can be compared by means of the two-sample K-S test. It should also be recalled that $x(t)$ in the interval $[0, 2T]$ can be tested for Gaussianity by means of the K-S goodness of fit test, with the mean and variance of the Gaussian population estimated by the sample mean and variance. Values for the two-sample K-S test ([13], p. 487) and for the K-S goodness of fit test with unknown mean and variance [19] at the 0.05 level of significance were used. After testing all 120 EEG segments of 64 sec duration for wide-sense stationarity and Gaussianity, each segment was subdivided into two segments of 32 sec duration which were also tested in the same manner. This procedure of successively subdividing and testing was repeated until all available EEG segments of 1 sec duration were tested. In total, $4M$ EEG segments of T sec duration were tested, $2M$ segments from the baseline ensemble and M segments from each of the anesthesia ensembles, where T and M are given by (7) and (8), respectively. For each of the three ensembles, the percentage of EEG segments of a specified duration which could be modeled as being wide-sense stationary, Gaussian, or both wide-sense stationary and Gaussian was calculated. All results were then corrected for type I errors arising from false rejection of the hypotheses being tested.

The computation of power spectra required as part of the previously described test for wide-sense stationarity was performed on an IBM 370/168 computer using the Direct Method, i.e., direct Fourier transformation of the data with consecutive averaging over frequency [7]. Each digitized EEG segment of T sec duration, consisting of a set of $128T$ sample values, was first tapered with a time window $W(t)$ of the form

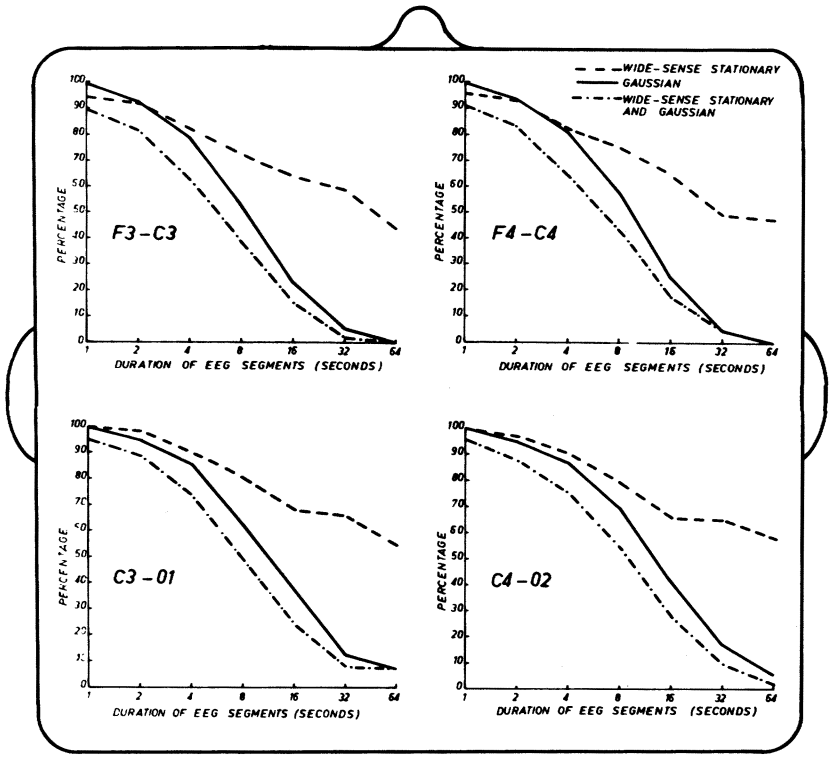

Fig. 4. Mean ensemble characteristics of the baseline EEG activity of 30 subjects who were resting with eyes closed. Results are based on a total of 3840 sec of EEG activity from each of 4 channels, collected in the form of two multichannel EEG samples of 64 sec duration per subject.

$$W(t)=\begin{cases}\frac{1}{2}\left(1-\cos\pi\frac{t}{0.1T}\right), & 0\leqslant t<0.1T\\ 1, & 0.1T\leqslant t<0.9T\\ \frac{1}{2}\left(1-\cos\pi\frac{T-t}{0.1T}\right), & 0.9T\leqslant t\leqslant T\end{cases}$$

and was then transformed via the Fast Fourier Transform algorithm. A periodogram was calculated from the complex Fourier coefficients for each fundamental frequency k/T Hz, where $k=0,1,\cdots,64T$. Smoothing of the periodogram was performed using a rectangular window with 7 non-zero coefficients. In this manner a set of $(64T+1)$ smoothed spectral estimates from 0-64 Hz was calculated for each EEG segment of T sec duration. The distribution function of the subset of spectral estimates between 1 Hz and 30 Hz was then used in the previously described test for wide-sense stationarity.

Results and Discussion

The results of our investigation are summarized graphically in Figs. 3-7. In Figs. 4-7, the results for each EEG channel are presented topologically, i.e., the results are located on a stylized representation of the head in a position corresponding to the location of the electrodes from which the EEG activity was recorded. Although all results have already been corrected for type I errors due to false rejections of the hypothesis being tested, type II errors due to false acceptances of the hypothesis may still exist. Also, these results are based on empirical tests for necessary, but not sufficient, properties that sample EEG segments must possess in order to be modeled as the output of a particular type of random process. For these reasons, the estimated percentages given in Figs. 3-7 therefore represent useful empirical upper bounds on the corresponding "true" percentages.

The effect of different sampling rates upon the outcome of statistical hypothesis tests is illustrated in Fig. 3. This marked and previously unexplored relationship may account for some discrepancies apparent in the literature. The problem arises from the assumption, made in the formulation of both the chi-square and the K-S tests, that the set of samples to be tested represents a set of independent random observations. In practice, as the rate of sampling a bandlimited EEG segment increases above the Nyquist rate, successive samples become more interdependent and the efficacy of statistical hypothesis tests is consequently affected. It is therefore not surprising that one study of 2 sec EEG segments which were sampled at 200 Hz concluded that resting EEG activity is Gaussian 66 percent of the time [5], while other studies of EEG segments of similar duration which were sampled at 5000 Hz concluded that resting EEG activity is strongly non-Gaussian [7]-[9]. Fig. 3 indicates that, if it is desired to investigate the characteristics of EEG segments by means of statistical hypothesis tests, the best tradeoff between the requirement to adequately sample a bandlimited signal and the desirability of satisfying the assumption of a statistically independent sample set is reached if the sampling rate is set as little above the Nyquist rate as is practicable.

The estimated statistical characteristics of the ensemble of baseline EEG activity are presented in Fig. 4. The percentage of EEG segments which can be modeled as being Gaussian, wide-sense stationary, or both is given for each of the 4 bipolar channels under consideration. In Fig. 4 the strong dependence of the results on the duration of the EEG segments being

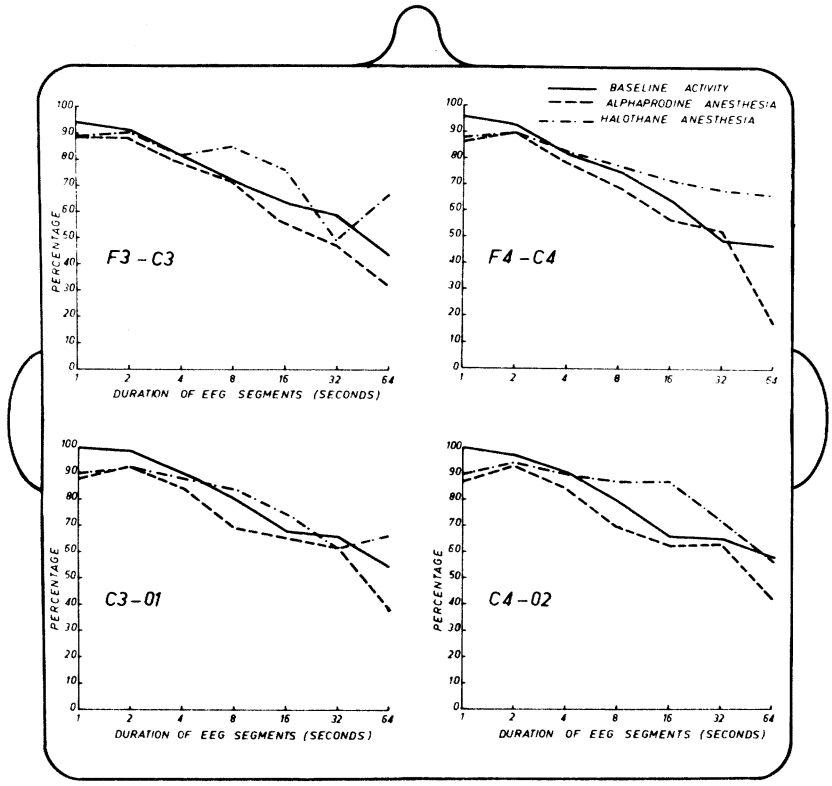

Fig. 5. Estimated percentage of EEG segments of various durations from each of the three different ensembles which can be modeled as wide-sense stationary.

tested is apparent. This dependence accounts for many of the discrepancies in the literature, e.g., our results are consistent with one previous finding [3] that two of four baseline EEG segments (of 8.33 sec duration) tested were Gaussian, and they are also consistent with another report that only 3.3 percent of 30 baseline EEG segments (of 52.8 sec duration) were found to be Gaussian [4]. The results in Fig. 4 also clearly differentiate between the properties of Gaussianity and stationarity; for example, in channel C4-02 over 57 percent of EEG segments of 64 sec duration were modeled as wide-sense stationary but only 5.3 percent were found to be Gaussian and less than 2.0 percent could be considered both Gaussian and wide-sense stationary. Fig. 4 also reveals striking similarities among corresponding results for all 4 channels, and even stronger similarities between results for pairs of bilaterally symmetric channels. Thus, while no obvious interhemispheric EEG differences were found, occipital EEG activity appears to be consistently more Gaussian and more stationary than frontal EEG activity.

In Figs. 5-7 the estimated statistical characteristics of baseline EEG activity are compared to the corresponding characteristics during alphaprodine anesthesia and during halothane anesthesia. The data base for each type of anesthesia consisted of 1920 sec of EEG activity from 15 subjects, i.e., two 64 sec segments per subject, and the baseline data consisted of a total of 3840 sec of EEG activity from all 30 subjects.

Fig. 5 shows the estimated percentages of sample EEG segments of various durations from each of the three different ensembles which can be modeled as wide-sense stationary. If the stationarity of EEG segments of the same duration is considered, it appears that EEG activity during halothane anesthesia is marginally more stationary than baseline activity, while EEG activity during alphaprodine anesthesia is slightly less stationary than baseline activity. The results in Fig. 5 indicate that, for sample EEG segments less than 32 sec in duration from any channel and from any of the three ensembles, the assumption of wide-sense stationarity may be valid more than 50 percent of the time.

Fig. 6 gives the estimated percentage of sample EEG segments from each ensemble which can be modeled as Gaussian. EEG segments from halothane anesthesia are generally less Gaussian than the baseline activity, particularly in channels C3-01 and C4-02, while EEG segments from alphaprodine anesthesia are marginally more Gaussian than the baseline activity. In all channels, EEG activity during halothane anesthesia is consistently less Gaussian than EEG activity during alphaprodine anesthesia.

In Fig. 7, the percentage of sample EEG segments from each of the three ensembles which can be modeled as both Gaussian and wide-sense stationary is presented. A bilateral symmetry is immediately apparent in these results. In all channels, the percentage of EEG segments from halothane anesthesia which are wide-sense stationary and Gaussian is markedly smaller than the corresponding percentage from alphaprodine anesthesia. Also, from Fig. 7 it is evident that less than 10 percent of the 64 sec EEG segments from any ensemble can be modeled as wide-sense stationary and Gaussian.

The estimated degree to which ensembles of EEG activity may be modeled as stationary and Gaussian, e.g., the results presented in Fig. 5 and Fig. 7, should be an important consideration in the choice of an appropriate technique for analyzing sample EEG segments from those ensembles. For example, the primary motivation for investigating the sta-

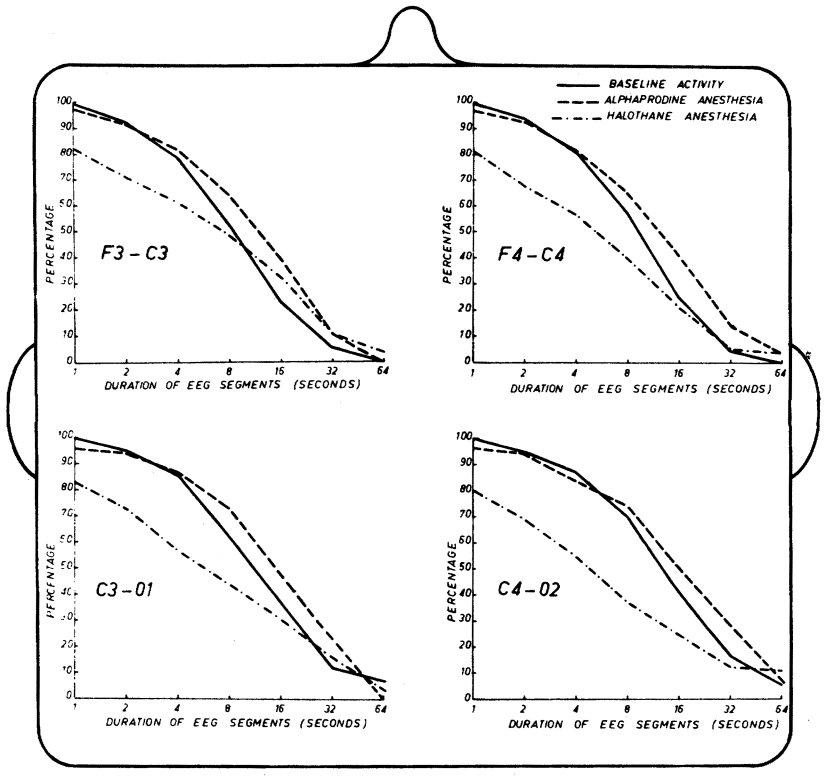

Fig. 6. Estimated percentage of EEG segments of various durations from the three different ensembles which can be modeled as Gaussian.

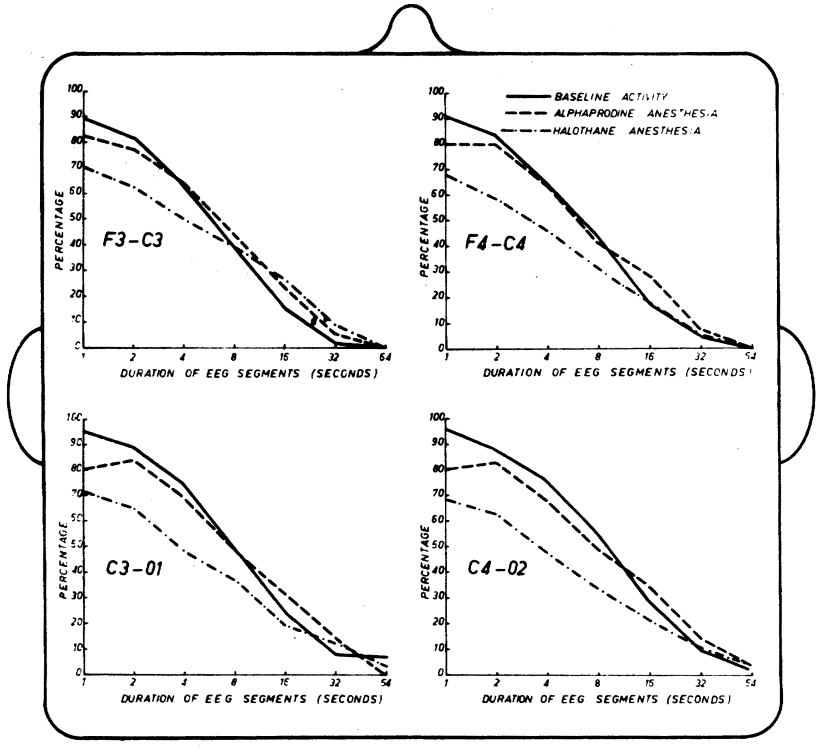

Fig. 7. Estimated percentage of EEG segments of various durations from the three different ensembles which can be modeled as both wide-sense stationary and Gaussian.

tistical characteristics of the three specific ensembles of EEG activity described in this paper was the expectation that the results would assist in the development of a computer-based system for monitoring the level of anesthesia during surgery by means of an automatic analysis of spontaneous EEG activity [17], [18]. In the development of such a system, decisions must be made with respect to the duration of the EEG segments to be analyzed, the rate at which the system should be updated, the choice of an analytic technique, and the significance which may be attached to the results of the analysis. It should be noted that the feasibility of employing EEG monitoring systems to continuously assess a patient's

status during sleep, serious illness, coma, and possible cerebral death is also currently being investigated by others, e.g., [20]-[22]. The statistical characteristics of the particular ensembles of EEG activity being analyzed in each instance should be an important consideration in the development of the monitoring system.

To illustrate how knowledge concerning the degree of stationarity of the three ensembles described in this paper might influence the development of a system for monitoring and analyzing EEG activity during anesthesia, we will briefly consider the problem of selecting the most appropriate duration for sample EEG segments on the basis of the results in Fig. 5. It would obviously be desirable to analyze EEG segments of long duration because the significance of any transient noise and artifact is thereby reduced, because a high resolution in the estimation of power spectra is possible, and because a potentially large data reduction can be realized if such segments can be adequately characterized. However, in the theoretical development of most analytic techniques the assumption is made that the signal under consideration represents a sample function from a random process that is at least stationary to some extent over the interval of interest. Fig. 5 indicates that the assumption of wide-sense stationarity for the three ensembles under consideration is only partially justified, even for EEG segments of relatively short duration. The *a priori* selection of the most suitable analytic technique therefore cannot be made on a firm theoretical basis. Under such conditions, the results in Fig. 5 indicate that the choice of 32 sec duration for sample EEG segments might, in this instance, represent a reasonable compromise. For all three ensembles at least one-half of the EEG segments of this duration could be modeled as wide-sense stationary. An analytic technique which assumes at least wide-sense stationarity could then reasonably be applied to the 32 sec segments and any inherent non-stationarity could be taken into account by some ancillary technique. For example, the previously described K-S D_2 statistics could be included in the analysis as parameters indicating the degree of non-stationarity of the segment being analyzed and hence could be used in interpreting the significance of the results. Alternatively, individual EEG segments could be tested for wide-sense stationarity as described previously and only those segments found to be stationary would be analyzed. If non-stationarities are to be considered for some particular EEG ensembles, and they cannot adequately be taken into account by such ancillary techniques, then a non-stationary analysis of the EEG could be attempted [23]-[25].

We now consider some implications of the results in Fig. 7 with respect to the choice of the most appropriate technique for analyzing EEG segments of a specified duration from any of the three ensembles. For the reasons stated previously, power spectrum analysis of the EEG segments would be preferable if the segments could be modeled as both wide-sense stationary and Gaussian. However, Fig. 7 shows that only a certain proportion of sample EEG segments may be so modeled, e.g., for all ensembles less than 50 percent of the 8 sec segments from any channel could be considered wide-sense stationary and Gaussian. It cannot therefore be assumed that spectral analysis will provide a sufficient characterization of such sample EEG segments. Where it is known that a certain proportion of the EEG segments to be analyzed cannot be modeled as the output of a stationary Gaussian random process, alternate analytic strategies might be considered. Of course, any analytic technique could arbitrarily be applied to the data in the hope that the results might somehow provide an *ad hoc* justification for its usage. However, if it can be assumed that most of the segments under consideration are wide-sense stationary, or that any inherent non-stationarity has been taken into account by one of the techniques described previously, then certain analytic strategies might be more profitably investigated. For example, if the EEG segments are stationary and only slightly non-Gaussian, ancillary parameters which indicate the degree of non-Gaussianity (e.g., skewness and kurtosis [26] or the previously described K-S D_1 statistic) might be employed in addition to spectral analysis. Alternatively, if the EEG segments to be analyzed are stationary but very non-Gaussian, then the information provided by EEG spectral analysis could be supplemented by the use of other analytic techniques, e.g., bispectral analysis [8].

Some further work is indicated. It has been suggested that, on the basis of the Central Limit Theorem, increased Gaussianity in observed EEG activity may reflect an increased degree of independence among individual cortical neural generators [5]. If one accepts this premise, then Fig. 6 and Fig. 7 indicate that the cortical generators are considerably more interdependent during halothane anesthesia than during alphaprodine anesthesia. The possible neurophysiological significance of this result could be investigated, perhaps by studies of EEG coherence in individual subjects and by considering more sample data from more channels. In addition, the technique described in this paper for estimating the degree of wide-sense stationarity and Gaussianity of an ensemble of EEG segments could obviously be applied to many other ensembles of EEG activity corresponding to other states of consciousness.

Acknowledgment

The authors would like to thank J. L. Berezowskyj, L. S. Root, and M. D. Low for the help which they provided in collecting and prescreening the EEG data. The assistance of M. E. Koombes during the digitization of the data is also gratefully acknowledged.

References

[1] K. S. Lion and D. F. Winter, "A method for the discrimination between signals and random noise of electrobiological potentials," *Electroenceph. clin. Neurophysiol.*, vol. 5, pp. 109-111, 1953.

[2] V. A. Kozhevnikov, "Some methods of automatic measurement of the electroencephalogram," *Electroenceph. clin. Neurophysiol.*, vol. 10, pp. 259-278, 1958.

[3] M. G. Saunders, "Amplitude probability density studies on alpha and alpha-like patterns," *Electroenceph. clin. Neurophysiol.*, vol. 15, pp. 761-767, 1963.

[4] J. Campbell *et al.*, "On the sufficiency of autocorrelation functions as EEG descriptors," *IEEE Trans. Bio-Med. Eng.*, vol. BME-14, pp. 49-52, 1967.

[5] R. Elul, "Gaussian behaviour of the electroencephalogram: changes during performance of mental task," *Science*, vol. 164, pp. 328-331, 1969.

[6] M. S. Weiss, "Non-Gaussian properties of the EEG during sleep," *Electroenceph. clin. Neurophysiol.*, vol. 34, pp. 200-202, 1973.

[7] G. Dumermuth *et al.*, "Numerical analysis of electroencephalographic data," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 404-411, 1970.

[8] G. Dumermuth *et al.*, "Analysis of the interrelations between

frequency bands of the EEG by means of the bispectrum," *Electroenceph. clin. Neurophysiol.*, vol. 31, pp. 137–148, 1971.
[9] E. Strehl, *Zur Amplitudenverteilung Des Spontanen Wach-Elektroencephalogramms*, Inaugural-Dissertation, Universität Zürich, Zürich, 1972.
[10] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1965.
[11] J. B. Thomas, *An Introduction to Statistical Communication Theory*. New York: Wiley, 1969.
[12] J. Bendat and A. Piersol, *Measurement and Analysis of Random Data*. New York: Wiley, 1966, pp. 88–91.
[13] B. W. Lindgren, *Statistical Theory*. New York: Macmillan, 1968, pp. 329–337.
[14] D. A. Darling, "The Kolmogorov-Smirnov, Cramer-von Mises tests," *Ann. Math. Stat.*, vol. 28, pp. 823–837, 1957.
[15] M. Kac, J. Kieffer and J. Wolfowitz, "On tests of normality and other tests of fit based on distance method," *Ann. Math. Stat.*, vol. 25, pp. 189–241, 1955.
[16] H. H. Jasper, "Report of the committee on methods of clinical examination in electroencephalography," *Electroenceph. clin. Neurophysiol.*, vol. 10, pp. 370–375, 1958.
[17] J. A. McEwen, G. B. Anderson, M. D. Low and L. C. Jenkins, "Monitoring the level of anesthesia by automatic analysis of spontaneous EEG activity," *IEEE Trans. Bio-Med. Eng.*, vol. BME-22, pp. 299–303, July 1975.
[18] J. A. McEwen, G. B. Anderson, M. D. Low, J. L. Berezowskyj and L. C. Jenkins, "Identification of anesthesia levels during surgery by computer-based EEG pattern recognition," in *Fifth Canadian Medical and Biological Engineering Conf. Dig.*, (Montreal, Can.), pp. 10.3a–10.3b, 1974.
[19] H. W. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *J. Amer. Stat. Ass.*, vol. 62, pp. 399–402, 1967.
[20] L. E. Larsen and D. O. Walter, "On automatic methods of sleep staging by EEG spectra," *Electroenceph. clin. Neurophysiol.*, vol. 28, pp. 459–467, 1970.
[21] B. R. Tharp, "The future of electroencephalography," in *Proc. Fifth Hawaii Int. Conf. on System Sciences–Computers in Biomedicine* (Honolulu, Hawaii), pp. 6–8, 1972.
[22] R. G. Bickford *et al.*, "A computer and enunciator system for estimation of brain death," in *Proc. San Diego Biomedical Symp.* (San Diego, Calif.), vol. 10, pp. 117–123, 1971.
[23] W. W. Wierville, "A new approach to the spectrum analysis of nonstationary signals," *IEEE Trans. Appl. Ind.*, vol. 82, pp. 322–327, Nov. 1963.
[24] P. D. Welch, "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 Electroacoust.*, vol. AU-15, pp. 70–73, June 1967.
[25] N. Kawabata, "A nonstationary analysis of the electroencephalogram," *IEEE Trans. Bio-Med. Eng.*, vol. BME-20, pp. 444–452, Nov. 1973.
[26] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics*, Part One. Princeton: Van Nostrand, 1954, pp. 99–103.

The Electrical Response of the Human Eye to Sinusoidal Light Stimulation

ARNE TROELSTRA, MEMBER, IEEE, AND CHARLES A. GARCIA

Abstract—**In order to extract maximum information from electroretinographic waveforms obtained under clinical conditions, it is necessary to have reliable quantitative methods to characterize both the amplitude and shape of these waveforms. When sinusoidally modulated light is used to stimulate the retina, the resulting ERG potentials are, in general, not sinusoidal due to the nonlinearities in the system. However, the responses are very reproducible and can easily be characterized by a few parameters based on a Fourier analysis. The more conventional flash ERG, although usually of higher amplitude, is much less reproducible in shape and needs many more parameters to characterize completely. Amplitude and phase characteristics can be understood on the basis of a simple model for the scotopic *B*-wave system and additive interaction by the photopic system. Changes in amplitude and phase characteristics with various experimental conditions could be predicted and were confirmed by subsequent experiments. The ultimate goal of this work is to improve the quantitative basis for clinical electroretinography and to provide the clinician with additional data which can be useful in the diagnosis of retinal diseases.**

Manuscript received April 22, 1974; revised January 9, 1975. This work was supported in part by National Institutes of Health Biomedical Sciences Support Grant 5 SO5 RR 07103.

A. Troelstra is with the Department of Electrical Engineering, Rice University, Houston, Tex.

C. A. Garcia is with the Department of Ophthalmology, University of Texas Medical School, Houston, Tex.

INTRODUCTION

THE electroretinogram (ERG) represents a change in electrical potential when the retina is stimulated with light. This electrical potential can be detected most easily using one electrode making contact with the cornea and another electrode connected to the temple or earlobe. Since the discovery of the electroretinogram about a hundred years ago, a tremendous amount of research has been directed toward an understanding of the significance of this electrical potential for the characterization of visual function and for use as an objective aid in clinical diagnosis. Presently, it is fairly well known at which retinal locations the various ERG components are generated [1], [2], [3], but the underlying mechanisms responsible for their generation remain to be established. It is possible to record ERGs from localized areas of the human retina [4], [5], [6] with properly selected combinations of stimulus and background intensities. In general, relatively high background and low stimulus intensities are most favorable, which makes these local responses low and only detectable with average response calculations involving large numbers of responses. Without these special precautions, the ERG reflects the overall retinal response to light stimulation, and the relation between