A Sequential Algorithm for Biological Event Detection Using Statistical Nonstationarity

John DiCecco, Jack Salisbury, PhD and Ying Sun, PhD,
Department of Electrical and Computer Engineering, University of Rhode Island, Kingston, RI 02881, USA

Abstract- **High dimension complex dynamical systems, such as those found in physiological processes, are often accompanied by nonstationarity. In many cases, the nonstationarity is caused by a physiologically significant event such as the prelude to ventricular fibrillation in cardiac arrest or the change of stasis by introducing pharmaceuticals. A need exists to be able to detect and monitor this change. Most conventional attempts at addressing this problem involve segmenting the time series and evaluating the statistics of the segments. The difficulty with this approach is that the nature of the nonstationarity can be transient, such that it is bounded by two, or more, regions of stationarity. Further, vacillation between stationary and nonstationary segments may continue for a significant portion of the time series. This paper will discuss the underlying statistical justification for asserting stationarity and the use of segmentation time series analysis techniques**.

***Index Terms*— Biomedical engineering, electrophysiology, neurophysiology, nonlinearity, stationarity, nonstationarity.**

I. Introduction

At any given moment, the communicatory action of the central nervous system is influenced by physical stimuli, chemical stimuli, and electrical stimuli. It is these nonlinear components that make analysis difficult, since it is problematic to account for the exact contribution of each underlying stimulus. Stochastic process analysis provides some measure of predictability, as stationarity is a function of the statistics of the time series. By establishing benchmark statistics from purely random independent identically distributed (iid) variables and comparing these to the studied time series, it is possible to detect and quantify the divergence. A time series will be arbitrarily defined as nonstationary if its mean changes by a threshold amount over the course of the signal. That is, a stationary signal is only defined as one where the statistics do not change over time, regardless of any arbitrary threshold [1]. A sequential algorithm has been developed which accurately determines when a test signal makes the transition from a stationary signal to a nonstationary one and then back to a stationary signal. Only current and previous values are used for calculation to allow for the transformation to a real-time algorithm.

John DiCecco is a doctoral student in the Department of Electrical and Computer Engineering at the University of Rhode Island, Kingston, RI 02881 USA (phone: 401-219-1021; e-mail: diceccoj@ ele.uri.edu).

Dr. Jack Salisbury is an alumnus of the University of Rhode Island and founder of Echo Technologies, Chepachet, Rhode Island, 02671, USA

Dr. Ying Sun is a professor of Biomedical and Electrical engineering at the University of Rhode Island, Kingston, RI 02881 USA

II. Methods

A three part control signal consisting of a Gaussian distributed random variable is generated to provide the benchmark test signal. The transition from the first segment of the signal to the final segment of the signal is along the slope of the difference between the two means. A protocol of 10% difference between the initial and final means and a 5% standard deviation of that difference is established (Fig. 1, top). A biological signal, which serves to provide the experimental performance of the algorithm, is obtained from the central nervous system (CNS) of the pond snail *Lymnaea stagnalis*. The administration of 10^{-6} Molar Acetylcholine (ACh) mimics a biological event by changing the neuronal membrane resting potential. ACh is administered using a gravity fed perfusion system at time 30s (sample 3000) and is removed by rinsing with saline at time 80s (sample 8000) (Fig. 2, top).

Dissection

The dissection of *L. stagnalis* begins with the administration of 0.36M MgCl to anesthetize the animal. The snail is removed from its shell and pinned to a Sylgard lined Petri dish in a solution of snail saline. Microscissors are used to expose the CNS, and the ganglia are removed intact and transferred to a smaller Sylgard lined Petri dish. The structure is stretched and pinned securely for insertion of the microelectrode [2].

Signal Acquisition and Segmentation

Recordings are made possible by the Gene Clamp 500 (Axon Instruments) and the PMD-1608FS Analog to Digital (A/D) converter (Measurement Computing).

The recorded signal is defined as an observed time series where the time t and index i are related by $t = t_0 + i\Delta t$, and the sample rate is given by $1/\Delta t$. The term *segment* refers to the portion of the time series that the sliding window is tested against. The segment should not be confused with the time series up to the window, as the segment is defined by the mode of the algorithm and thus has a length established by the mode. That is, when the algorithm detects a transition point, it determines the nature of the data based on its slope, and defines that point as the beginning of the next segment.

In this analysis, a time series is defined as stationary if the mean remains constant over the range of the function with a slope statistically equal to zero. This is, of course, a loose definition of stationarity, but it is sufficient for testing purposes. For detection purposes, a statistically significant difference in mean is one that exceeds 0.95, or 95%

confidence, in the statistics between the segment and the window [1]. Similarly, a statistically significant difference between the calculated slope and zero will have a 0.02 to 0.05 threshold. This range is necessary at this stage to compensate for the noise in the experimental data.

III. Calculations

The most rigid definition of stationarity is established for a random process in which the statistics do not change over the entire length of the time series. Without *a priori* knowledge of the length of the time series or the disposition towards nonstationarity, the assessment of the critical statistics must involve the windowed statistics versus the statistics of the segment. The window is compared to the threshold established by the segment and any statistically significant deviation is detected. The slope of the windowed portion of the signal is determined by

$$slope = \frac{\sum x_i y_i - (\sum x_i \sum y_i)/n}{\sum x_i^2 - (\sum x_i)^2 / n} \quad (1)$$

where x_i and y_i represent the x and y components of the time series and n is the window length in samples [3]. Statistically significant changes between the window mean and the segment mean are determined by using the paired-t test

$$paired\ t = \frac{\mu_{seg} - \mu_{win}}{\sqrt{\left(\frac{n_{seg} S_{seg}^2 + n_{win} S_{win}^2}{n_{seg} + n_{win}}\right)\left(\frac{1}{n_{seg}} + \frac{1}{n_{win}}\right)}} \quad (2)$$

where μ_{seg} is the mean of the segment up to and including the window and μ_{win} is the mean of the window [3]. The *S* term is the sampled variance and the n_{seg} and n_{win} terms are the segment size and window size in samples, respectively.

IV. Results

Test 1 has mean equal to 2.2 for the first 1000 samples and mean equal to 2 for the final 1000 samples (Fig. 1, A). The algorithm successfully determines the boundary points of the nonstationarity by evaluating the statistical significance of the fluctuations between the mean of the segment and the mean of the window (Fig. 1, B).

The performance of the algorithm on Test 2 is shown in Fig. 2. While it successfully identified the addition of the ACh at just after 30s and the return to a stationary mean at 80s, the return values were not as accurate as for the control data. This is improved by fine tuning the variable parameters such as window size and statistical sensitivity.

V. Discussion

A sequential algorithm for biological event detection using statistical nonstationarity has been proposed. The algorithm uses only causal data and is capable of being implemented in real time. The nature of the statistical evaluation using only causal samples renders the algorithm useful for linear as well as nonlinear data sets [4]. The performance of the algorithm is heavily reliant on chosen parameters such as window size.

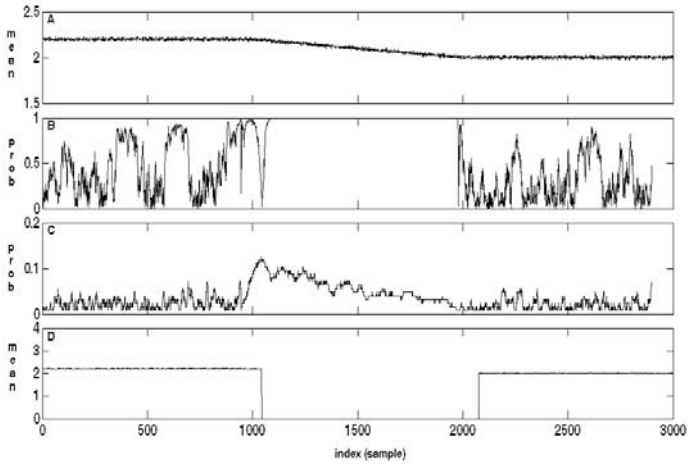

Figure 1. From top: A) the time series Test 1 - synthetic data, B) the t-test for statistically significant difference between the segment mean and the window mean reveals both the change at 1000 and the change at 2000, C) the time series as a function of its slope compared to zero. (Note that the transition period from 1000 to 2000 reveals that the slope is not zero, as expected), D) the data series after the slope of the transition has been subtracted from the time series.

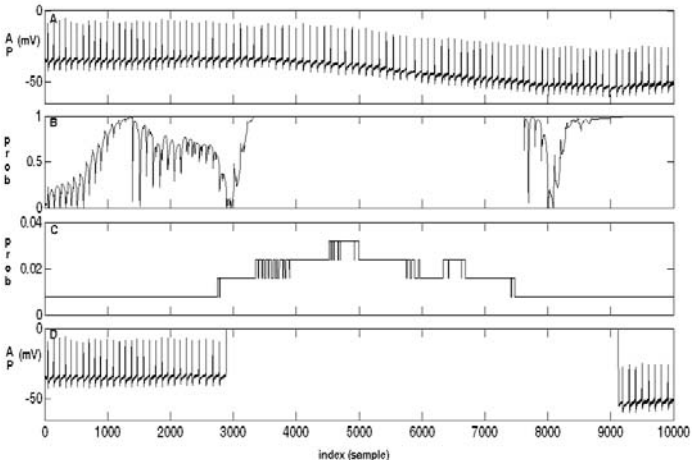

Figure 2. From top: A) the time series Test 2 - experimental data, B) the t-test for statistically significant difference between the segment mean and the window mean reveals both the change at 3000 and the change at 8000 (30s and 80s), C) the time series as a function of its slope compared to zero. (Note that the transition period from 3000 to 8000 reveals that the slope is not zero, as expected), D) the data series after subtracting the transition from the time series. AP - action potential, prob – probability.

There is also the issue of arbitrarily designating the detection criteria. The addition of dimensional analysis will aid in determining accurate parameter designation. This will be addressed in future revisions of the algorithm.

Acknowledgement

This study was supported in part by the URI Partnership in Physiological Measurements and Computing.

References

[1] Leon-Garcia, Alberto. *Probability and Random Processes for the Electrical Engineer*. Addison-Wesley Publishing, 1994.

[2] Borradaile, L.A. and Potts, F.A. The Invertebrata: A manual for the use of students. Cambridge University Press, 1958.

[3] Bahn, Anita K. *Basic Medical Statistics*. Grune & Stratton, Inc. 1972

[4] Sprott, Julian. *Chaos and Time-Series Analysis*. Oxford University Press, 2003.