ELSEVIER

BioSystems 40 (1997) 111–118

BioSystems

Spike train patterning and forecastability

André Longtin*, Daniel M. Racicot

Département de Physique, Université d'Ottawa, 150 Louis Pasteur, Ottawa, Ontario, Canada K1N 6N5

Abstract

Theories of neural coding rely on a knowledge of correlations between firing events. These correlations are also useful to validate biophysical models for the neural activity. We present a methodology for validating models based on the assessment of linear and non-linear correlations between variables derived from the spike train. The firing pattern of an electroreceptor is analyzed in this framework. We show that a purely stochastic model fails to capture the essential correlations between interspike intervals, even though it reproduces the interval histogram and certain spike train spectral features. However, a biophysical model, based on the Fitzhugh-Nagumo equations with noise, does exhibit many of the correlations seen in the data, including those between successive firing phases.

Keywords: Neural modeling; Point processes; Non-linear dynamics; Phase locking; Noise; Forecasting

1. Introduction

Neurons communicate mostly using propagated action potentials. The codes used for this purpose are not well known. To properly constrain theories of neural coding and information processing, it is important to characterize interspike interval (ISI) patterning and correlations. These are functions of the inputs to the neuron and of the non-linear dynamics of its conductances. A knowledge of these dynamics provides insight into the nature of ISI correlations. In turn, the nature of ISI correlations will influence the extent to which the timing of spikes determines firing in postsynaptic neurons. Correlations between dynamical variables governing neural activity also imply a certain degree of forecastability. Alternately, measures of forecastability can provide insight into the origin of ISI correlations (Longtin, 1993a). This forecastability, i.e. the function relating future ISI's to past ISI's, can be linear or non-linear.

The goal of this paper is to present a methodology for validating biophysical models of neural activity using the assessment of linear and non-linear correlations between ISI's. Our working hypothesis is that a good model for an ISI sequence should reproduce its main correlations (Longtin, 1993a). Our study builds on recent work (Longtin and Racicot, 1996) which contrasts linear and non-linear ISI analyses, and makes careful use of 'surrogate data' techniques (Theiler et al. 1991) in the context of point processes. Surro-

* Corresponding author.

PII S0303-2647(96)01636-X

gate data sets preserve some features of the raw data, such as the interval histogram (ISIH) or autocorrelation or both, but are otherwise random. In the context of spike trains, surrogate data techniques allow one to ascribe measured correlations to specific features of the data. Here we build simple models for the firing activity and generate spike trains from them. These serve, in some sense, as 'biophysical' surrogate data sets.

We illustrate our approach by analyzing spike trains from primary afferents of electroreceptors. These exhibit aperiodic phase locking or 'skipping', a firing pattern seen in many sensory neurons (Longtin, 1995). Our study aims to determine whether stochastic forces are present (Holden, 1976) and whether ISI correlations decay quickly. This knowledge will increase our understanding of the biophysics of the receptors and axonal spike generating zones, and of how physical stimuli modify the autocorrelation and cross-correlation of afferent ISI's. The methods of assessment of ISI correlations are reviewed in Section 2, and are applied to electroreceptor data in Section 3. We then apply the same correlation analyses to a stochastic model in Section 4 and to a biophysical model in Section 5. The comparison between firing phase correlations for the data and the models is also presented in Section 5, and the paper concludes in Section 6.

2. Linear or non-linear correlations?

Correlations in neural point processes can be studied using either the spike train, seen as a sequence of delta functions at the firing times, or the ISI sequence derived from the spike train (Moore et al. 1966). The ISIH is independent of the temporal properties of the spike train. Temporal information is revealed by autocorrelation functions and power spectra. The absence of temporal correlations between firing times produces flat spike train and interval power spectra. Deviations from flatness imply correlations, which can be used to forecast future ISI's, using for example an auto-regressive moving average (ARMA) model. Non-linear forecastability goes beyond such simple linear stochastic models by revealing non-linear relationships between successive ISI's. Further, it is a characteristic of the non-linear dynamics of firing, i.e. a dynamical invariant (Farmer and Sidorowich 1987).

Here the nearest-neighbor non-linear forecasting method described in Sauer (1994) is used. A sequence of N ISI's $(I_1, I_2, \ldots, I_N)$ allows us to construct an m-dimensional embedding $\vec{x}_n = (I_{n-m+1}, \ldots, I_{n-1}, I_n)$. The idea behind non-linear forecasting is to find a function F which maps a neighborhood of each point on an ISI 'trajectory' into the image of that neighborhood in the embedding space: $\vec{x}_{n+1} = F(I_{n-m+1}, \ldots, I_{n-1}, I_n)$. We will be interested in predicting only the next ISI, I_{n+1}, using the previous m ISI's. The predicted interval f is the average of the I_{n+1}'s over all nearest neighbors in the embedding space, which correspond to similar sequences of m successive ISI's. We impose that neighborhood sizes comprise one percent of the total number of ISI's. A normalized prediction error (NPE) is computed by first averaging the squares of all the prediction errors $(f-I)^2$ over the whole data set, and then dividing this result by the average error incurred by predicting the mean of the ISI sequence, $\bar{I}$:

$$\mathrm{NPE} = \left[\frac{\langle (f-I)^2 \rangle}{\langle (\bar{I}-I)^2 \rangle} \right]^{1/2}. \qquad (1)$$

We will compute the NPE for $m = 1$–8. An NPE value close to zero signifies that there is predictability; a value near one signifies there is little predictability, and one may as well forecast the mean $\bar{I}$. Linear correlations can fool non-linear forecasting algorithms by giving NPE values less than one (Tsonis and Elsner, 1992). It is important, therefore, to also compute the NPE for surrogate data sets.

3. Data from electroreceptor afferents

We illustrate our approach using a spike train from a primary afferent of an electroreceptor of the weakly electric fish *Apteronotus leptorhynchus*. The fish produces an AC electric field (called EOD), the modulations of which are used to locate food and communicate with other fish.

This rhythm drives the electroreceptors on the skin of the fish. Action potentials are recorded from the afferent axon from this receptor in the absence of EOD modulations. The receptor is thus driven by an almost perfectly periodic electric field. Under these conditions, the afferent axon fires near a preferred phase of this stimulus, but 'skips' an apparently random integer number of stimulus cycles between firings.

The first 60 ISI's are shown in Fig. 1A. The ISIH, with peaks at integer multiples of the EOD driving period (1.51 ms), is shown in Fig. 1B. The return map of successive ISI's in Fig. 1C reveals clusters of points having a slight stretch in a direction of negative slope, a signature of phase locking (Longtin and Racicot, 1996). The autocorrelation of successive ISI's in Fig. 1D decays to zero within two lags. The negative correlation at lag 1 suggests an alternation between long and short ISI's, i.e. between long and short skips. The NPE for the raw data is plotted in Fig. 2. It is significantly less than one for all m. It also differs from the NPE for surrogates obtained by shuffling the raw ISI sequence. It is only slightly lower than the NPE for surrogates having the same ISIH and autocorrelation of intervals

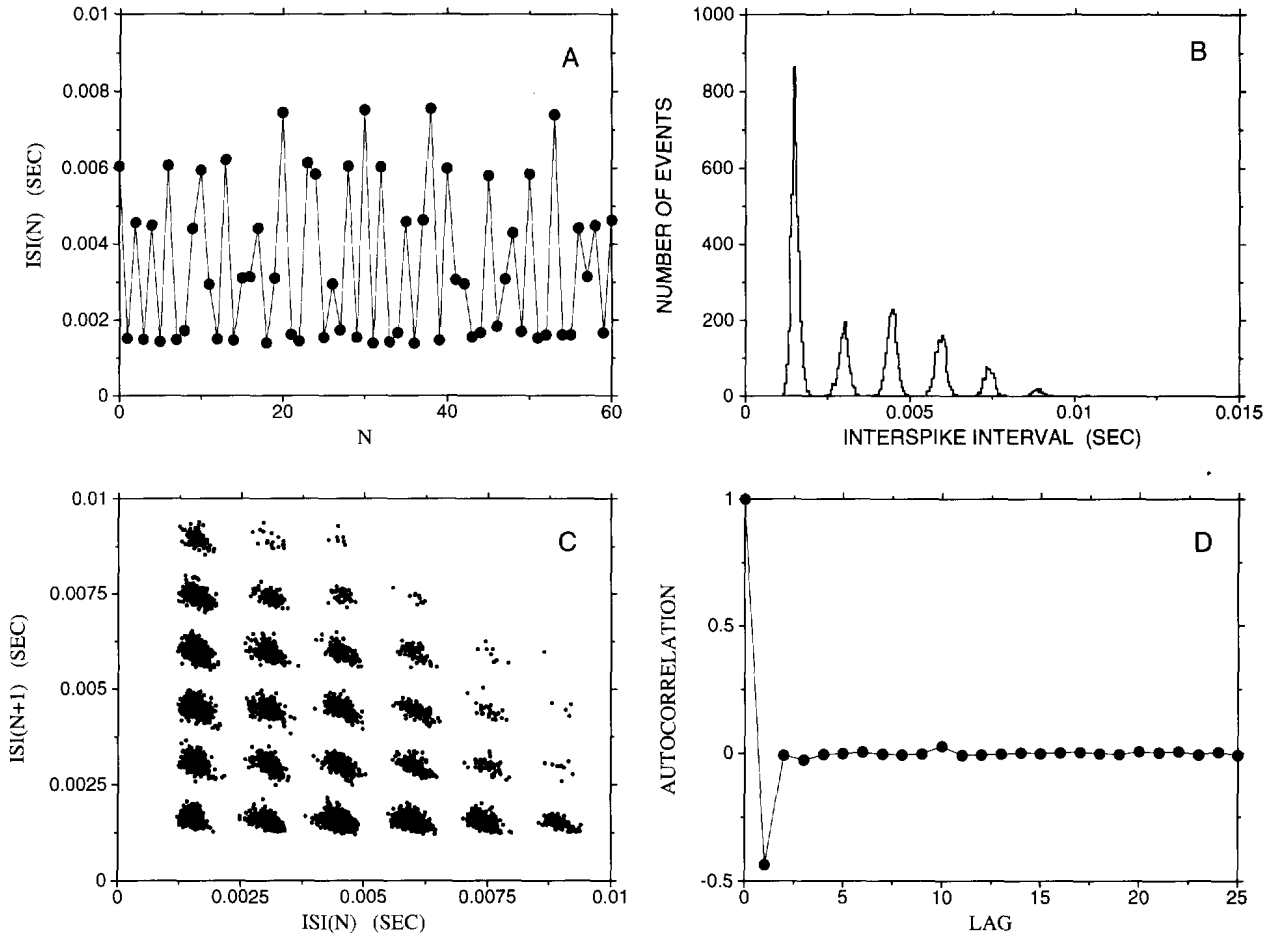

Fig. 1. (A) Interspike interval (ISI) versus interval number from extracellular recordings of a primary afferent neuron of a weakly electric fish (data provided courtesy of Joseph Bastian, U. Oklahoma). (B) ISIH for the data shown in (A). The 9165 intervals are discretized in 300 bins between 0 and 15 ms. (C) Return map of 9165 successive ISI's. (D) Autocorrelation of successive ISI's estimated using 50 lags.

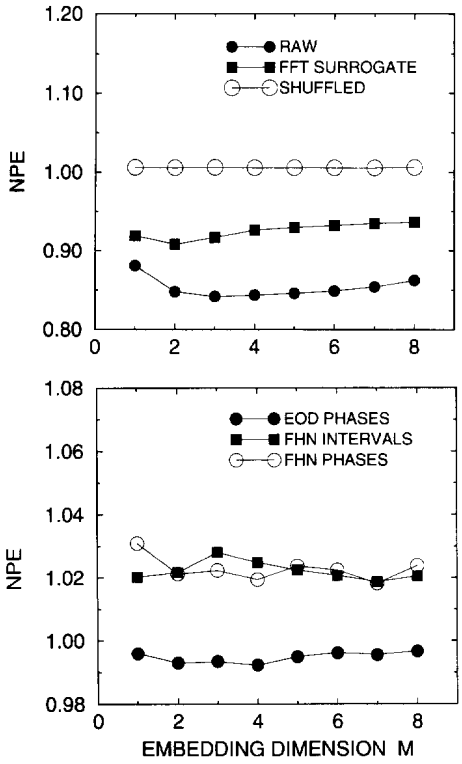

Fig. 2. Upper panel: Normalized prediction error (NPE) versus embedding dimension m for the data in Fig. 1. The NPE is shown for the raw ISI sequence, and for an average of 10 shuffled surrogates and 10 amplitude-adjusted phase-randomized surrogates. The two-σ error bars for the surrogates are smaller than the size of the symbols used. Lower panel: NPE versus m for the neuron firing phases, along with the NPE for the intervals and phases from the FHN model (Eqs. (4–6)).

('amplitude-adjusted' phase-randomized surrogates: see Theiler et al., 1991), suggesting a small amount of non-linear predictability (Longtin and Racicot, 1996).

4. A stochastic model with same ISIH

We can gain insight into the neural dynamics underlying this ISI data by considering a stochastic process that mimicks its ISIH. The process generates random ISI's, each interval I being the sum of a discrete Poisson variable I_P and a Gaussian variable I_G. I_P represents the number of skipped stimulus cycles between two spikes. It labels the ISIH peak number and determines its height. The I_P distribution is not necessarily monotonic. Its single parameter can be adjusted such that the second peak is larger than the first, as is sometimes seen in biological data. I_G accounts for the firing phase jitter, which does not have any cycle-to-cycle correlations (Section 5.2). The variance σ^2 of I_G determines the width of individual ISIH peaks. The stochastic process I, which could embody a biophysical model for e.g. the ISIH peak shapes and heights, does not have the linear ISI correlations of the raw data, since by construction I_G and I_P have zero autocorrelation and cross-correlation. Consequently, the probability distribution $P(I)$ of intervals is the convolution of the Gaussian and discrete Poisson distributions: $P(I) = P_P(I_P) * P_G(I_G)$. If I_P follows the distribution

$$P_P(I_P) = e^{-a} \sum_{k=1}^{\infty} \frac{a^k}{k!} \delta(I_P - kT_0), \tag{2}$$

where T_0 is the EOD period, the resulting normalized distribution for I is

$$P(I) = \left(\frac{2}{\pi\sigma^2}\right)^{1/2} \left[\sum_{n=1}^{\infty} \frac{a^n}{n!} \operatorname{erfc}\left(\frac{-nT_0}{\sqrt{2\sigma^2}}\right)\right]^{-1} \sum_{k=1}^{\infty} \frac{a^k}{k!} e^{-(I-kT_0)^2/2\sigma^2} \tag{3}$$

The sum starts at $k = 1$, i.e. the first peak at the origin is discarded. We adjust σ^2 and a (the mean of I_P) to obtain an ISIH similar to that in Fig. 1B. The resulting ISIH is shown in Fig. 3A, in which the ISI's are in units of T_0. A sequence of 2048 ISI's is generated using these parameters. The return map for this sequence is shown in Fig. 3B. The power spectrum of the associated spike train (not of the ISI sequence!) is shown in Fig. 3C. This spectrum is similar to that in Fig. 3D for a shuffled version of the raw data. Thus, this stochastic model has similar phase locking properties as the shuffled neuron data. The peaks in the spectrum of the raw spike train (not shown) are higher and sharper than those in Figs. 3C–D since, in the raw data, the firings keep in step with the stimulus, resulting in negatively-sloped clusters (Longtin and Racicot 1996). The shuffling, as

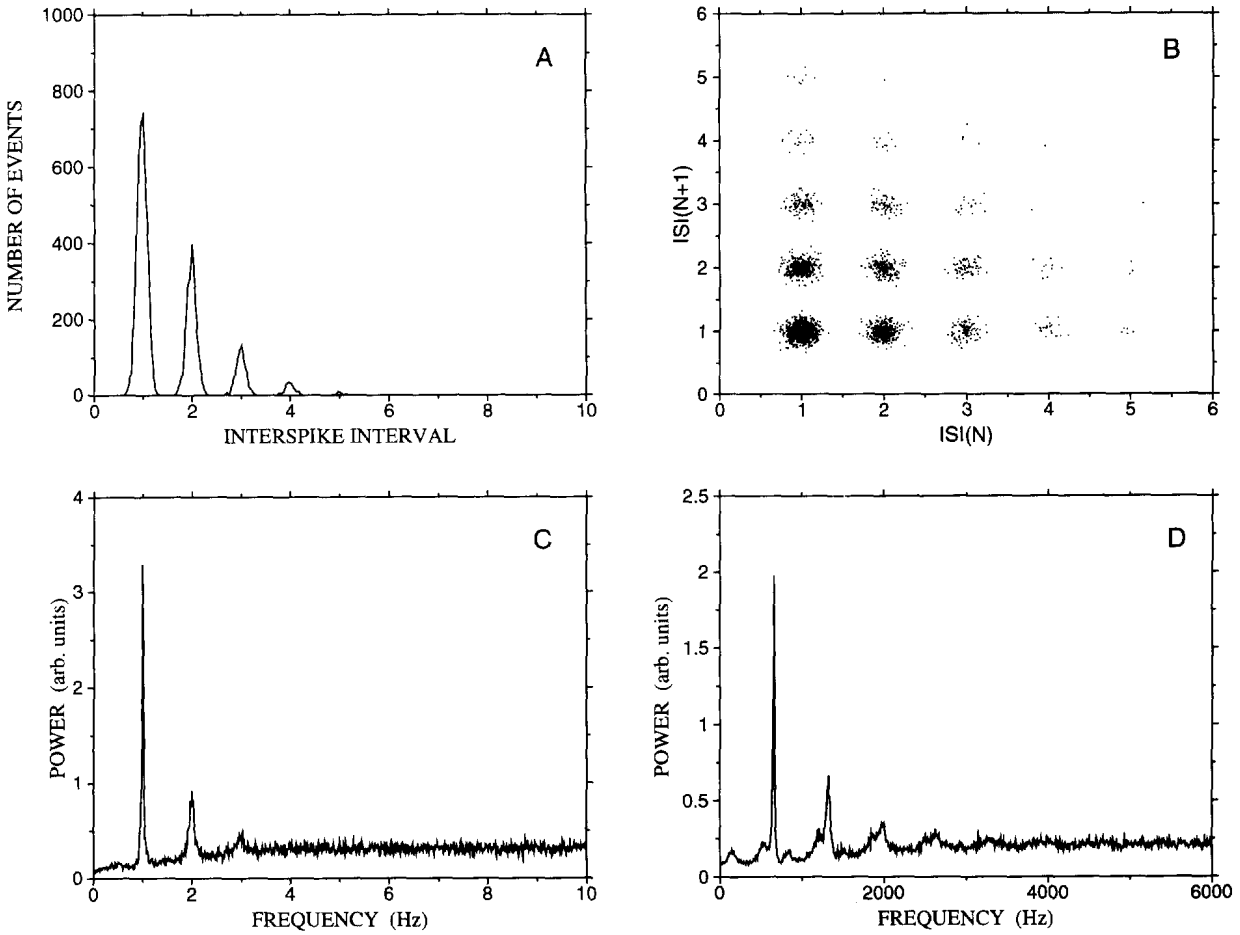

Fig. 3. Firing statistics of a stochastic model for the data in Fig. 1. I_P has period $T_0 = 1$ and mean $a = 1.0$. I_G has zero mean and $\sigma = 0.1$. (A) ISIH constructed from 2048 intervals. (B) Return map of successive ISI's. (C) Power spectrum of the spike train constructed by averaging 50 spectra from as many realizations of the stochastic model. Each spectrum was calculated from a 204.8 s data set. (D) Spike train power spectrum from a randomly shuffled version of the raw ISI sequence in Fig. 1, obtained by averaging 103 spectra, each using a 0.303 s window.

well as this model, produce clusters that are more symmetric (compare Fig. 1C to Fig. 3B). Not surprisingly, this stochastic model has an NPE equal to one.

5. A biophysical model

5.1. Noise-induced skipping

We next consider a more realistic model for action potential generation, the Fitzhugh-Nagumo equations (FHN) with periodic and stochastic forcing (Longtin, 1993b):

$$\epsilon \frac{dv}{dt} = v(v-a)(1-v) - w + \eta(t) \quad (4)$$

$$\frac{dw}{dt} = v - dw - [b + r \sin(\beta t)] \quad (5)$$

$$\frac{d\eta}{dt} = -\lambda\eta + \lambda\xi(t). \quad (6)$$

The periodic forcing is added to the recovery variable w, in order to match previous studies of forcing without noise (see Alexander et al., 1990). For frequencies higher than those used here, the

forcing should instead be added to v. Since $\xi(t)$ is zero-mean Gaussian white noise with $\langle \xi(t)\xi(s)\rangle = 2D\delta(t-s)$, $\eta(t)$ is an Ornstein-Uhlenbeck process, characterized by its correlation time $t_c = \lambda^{-1}$ and its variance $D\lambda$. The parameters, which are not meant to match the experiment precisely, are such that in the absence of periodic forcing and noise, the system is excitable. The separation between the fixed point and the threshold is set by b; as b increases with $r = 0$, a Hopf bifurcation to periodic firing occurs, with period 0.77. For the parameters chosen, this time scale has a negligible influence on ISI correlations. Also, the periodic forcing alone can not induce spikes.

The parameters D (noise intensity) and the amplitude r can be adjusted until an ISIH similar to that in Fig. 1B is obtained (Fig. 4A). This regime can not produce a second peak smaller than the first and third peaks as in Fig. 1B, a feature which is atypical for simple noise-induced firing (Longtin, 1995), and which may arise because other behaviors are sampled by the noise in the electroreceptor. This model also exhibits the slanted clusters in the return map (Fig. 4B), a consequence of the genuine phase locked nature of these dynamics. This implies that a 'deviate' variable (Longtin, 1993a), measuring which side of an ISIH peak an ISI falls, should exhibit a negative correlation (Longtin and Racicot, 1996). The FHN data indeed exhibit this correlation, as well as a sharp spike train power spectrum similar to that of the raw data (both not shown). However, the FHN data has a flat ISI autocorrelation, and thus differs in this respect from the raw data (Fig. 1D). Finally, the NPE of ISI's from this model is close to one (Fig. 2), again different from the value of ≈ 0.85 obtained for the raw data. This is due in large part to the flat ISI autocorrelation.

5.2. Forecasting the firing phases

We now investigate correlations between successive phases at which the primary afferent and the FHN model fire. For the FHN model, the phase measures the time between a spike and the preceding maximum of the sinusoidal stimulus.

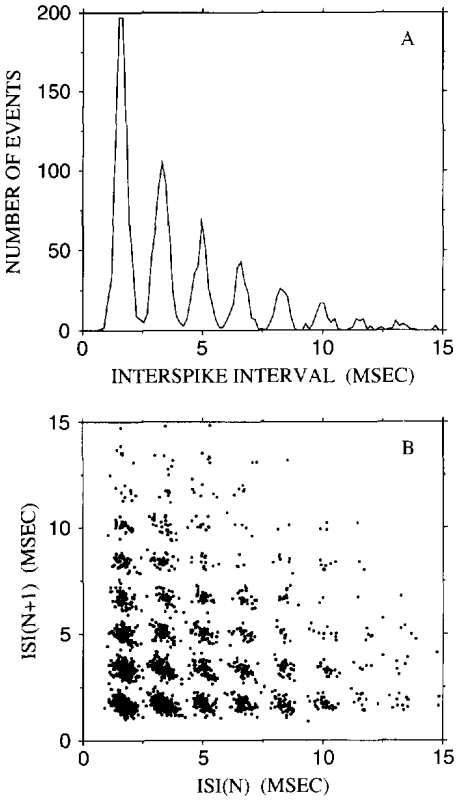

Fig. 4. ISIH (A) and ISI return map (B) for the Fitzhugh-Nagumo model (Eqs. 4–6) of the data in Fig. 1. Parameters are $a = 0.5$, $b = 0.12$, $d = 1$, $r = 0.06$, $\epsilon = 0.005$, $\beta = 3.75$, $D = 6 \times 10^{-6}$, and $\lambda = 100$. A fixed step (0.0025) fourth order Runge-Kutta method was used to generate 2048 ISI's. The spiking threshold was set at 0.5, and an absolute refractory period of 0.25 was used to reject false spikes.

For the EOD data, the phase is the time between a spike and the zero-crossing of the EOD stimulus immediately preceding it. The return map of successive phases is shown in Fig. 5A for the FHN model and Fig. 5B for the raw data. Both phases have insignificant serial correlation, linear (not shown) or non-linear (NPE ≈ 1 — see Fig. 2). This implies that the firing probability versus phase in a given cycle does not vary from cycle to cycle. The FHN model thus captures this feature of the raw data. Note that it is possible (in the case of the raw data) for successive ISI's to be negatively correlated in the absence of phase correlations. This requires a detailed explanation that will be given elsewhere. Finally, the EOD

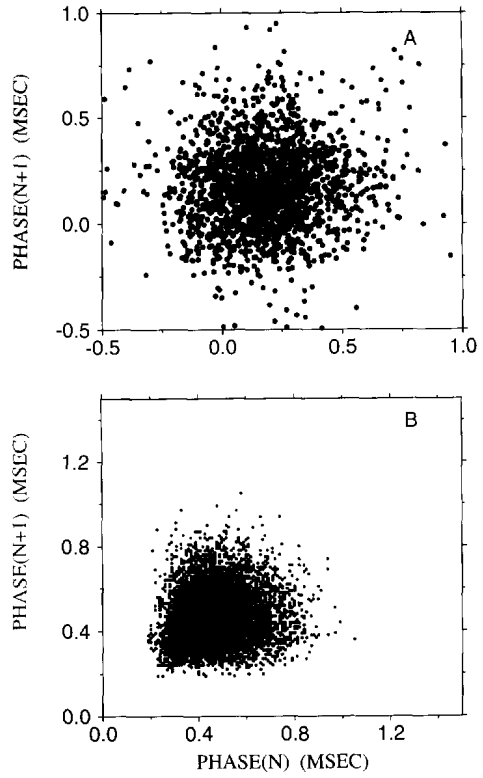

Fig. 5. Return map of firing phases for the Fitzhugh-Nagumo model (A) and for the raw data in Fig. 1 (B).

period was found to fluctuate between 1.46 and 1.56 ms, with a negative correlation at lag one (not shown). Although ten times smaller than the fluctuations in phase, these fluctuations may affect ISI correlations, and account for some of their non-linear forecastability.

6. Conclusion

Our proposed approach to studying spike train generation and patterning is based on the comparison of analyses of ISIH's and ISI correlation on data from neurons and from biophysical models of these neurons. In the context of surrogate data-based assessments of non-linearity (Theiler et al., 1992), these models serve as 'biophysical surrogates', and correspond to more refined 'null hypotheses' about the firing dynamics. Specifically, we have found that a purely stochastic model can reproduce the ISIH and certain features of the ISI return map and spike train power spectrum from an electroreceptor primary afferent. A model for noise-induced skipping, based on the Fitzhugh-Nagumo equations, exhibits even more features of the raw data by also producing phase locked behavior (slanted clusters in the return map) and uncorrelated firing phases. However, this model did not show negative correlation between successive ISI's and an NPE value below one, in contrast with the raw data. Possible reasons for this are that the receptor and/or afferent is dynamically more complex than Eqs. (4–6) (the power spectra difference is similar to that between Figs. 3C and 3D). Also, Eqs. (4–6) may not adequately model the refractory behavior at the frequencies used. The issue of whether ISI correlations have any functional significance for this fish can only be resolved by recording postsynaptically. Meanwhile, our approach will be useful to validate biophysical models for the transduction process, and already points to the importance of noise in this process.

Acknowledgements

This research was supported by NSERC (Canada) and by NIMH (USA) grant 1-R01-MH47184-01. The authors are grateful to Joe Bastian, for providing the physiological data, and to Leonard Maler for useful comments on this work.

References

Alexander, J.C., Doedel, E.J. and Othmer, H.G., 1990, On the resonance structure in a forced excitable system. SIAM J. Appl. Math. 50, 1373–1418.

Farmer, J.D. and Sidorowich, J.J., 1987, Predicting chaotic time series. Phys. Rev. Lett. 59, 845–848.

Holden, A.V., 1976, Models of the stochastic activity of neurones, Lect. Notes in Biomath. Vol. 12 (Springer Verlag, Berlin).

Longtin, A., 1993a, Nonlinear forecasting of spike trains from sensory neurons. Int. J. Bif. Chaos 3, 651–661.

Longtin, A., 1993b, Stochastic resonance in neuron models. J. Stat. Phys. 70, 309–327.

Longtin, A., 1995, Mechanisms of stochastic phase locking. Chaos 5, 209–215.

Longtin, A. and Racicot, D.M., 1996, Assessment of linear and non-linear correlations between neural firing events, Proc. of the Nonlinear Dynamics and Time Series Conf.,

Montréal, July 1995. Fields Inst. Commun., Am. Math. Soc. (in press).

Moore, G.P., Perkel, D.H. and Segundo, J.P., 1966, Statistical analysis and functional interpretation of neuronal spike data. Annu. Rev. Physiol. 28, 493–522.

Sauer, T., 1994, Reconstruction of dynamical systems from interspike intervals. Phys. Rev. Lett. 72, 3811–3814.

Theiler, J., Galdrikian, B., Longtin, A., Eubank, S. and Farmer, J.D., 1991, Using surrogate data to detect non-linearity in time series, in: Nonlinear Prediction and Modeling, M. Casdagli and S. Eubank (eds.) (Addison-Wesley, Redwood City, Ca) pp. 163–188.

Tsonis, A.A. and Elsner, J.B., 1992, Nonlinear prediction as a way of distinguishing chaos from random fractal sequences. Nature 358, 217–220.