Annals of Biomedical Engineering, Vol. 27, pp. 581–591, 1999
Printed in the USA.

0090-6964/99/27(5)/581/11/$15.00

Application of a Novel Modeling Method to the Nonstationary Properties of Potentiation in the Rabbit Hippocampus

M. IATROU, T. W. BERGER, and V. Z. MARMARELIS

Department of Biomedical Engineering, University of Southern California, Los Angeles, CA

(Received 16 October 1998; accepted 30 June 1999)

Abstract—This paper presents the first application of a novel methodology for nonstationary nonlinear modeling to neurobiological data consisting of extracellular population field potentials recorded from the dendritic layer of the dentate gyrus of the rabbit hippocampus under conditions of stimulus-induced potentiation. The experimental stimulus was a Poisson random sequence with a mean rate of 5 impulses/s applied to the perforant path, which was sufficient to induce a progressive potentiation of perforant path-evoked granule cell response. The modeling method utilizes a novel artificial neural network architecture, which is based on the general time-varying Volterra model. The artificial neural network is composed of parallel subnets of three-layer perceptrons with polynomial activation functions, with the output of each subnet modulated by an appropriate time function that models the system nonstationarities and gives the summative output its time-varying characteristics. For the specific application presented herein these time functions are sigmoidal functions with trainable slopes and inflection points. A possible mapping between the nonstationary components of the model and the mechanisms underlying potentiation changes in the hippocampus is discussed. © *1999 Biomedical Engineering Society.* [S0090-6964(99)00305-7]

Keywords—Nonstationary nonlinear modeling, Volterra models, Time-varying artificial neural network, Dentate gyrus, Potentiation, Hippocampus.

INTRODUCTION

One of the most challenging issues in the mathematical modeling of neurobiological systems from stimulus-response data has been the occasional presence of nonstationarities in the functional characteristics of some neurobiological processes. This challenge is further augmented when these systems exhibit nonlinear dynamics. Although several general nonlinear modeling methodologies have been proposed using Volterra models,[11–14] nonlinear auto-regressive moving average with exogenous variable models[4,10] or artificial neural networks,[4,5,15–17,19] the development of modeling for nonstationary systems has been lagging. The nonstationary problem has been dealt with either in a piecewise quasistationary fashion or by means of adaptive recursive algorithms that track the time-varying parameters of the system,[1] but do not provide models of global validity. Models of global validity can be obtained by a method based on ensemble averaging[11] and by a nonparametric temporal expansion method when the input is white noise.[11,12] The method proposed herein removes this restrictive input requirement.

The proposed methodology for the modeling of nonstationary and nonlinear systems was introduced in Ref. 7 and is suitable for systems with finite memory receiving broadband stimulus signals. This approach employs a novel artificial neural network (ANN) architecture, which is based on the theory of functional expansions of nonlinear, nonstationary systems (time-varying Volterra series),[13] and tapped-delay feed-forward ANNs with a single hidden layer and polynomial activation functions.[15] The resulting architecture contains a specific structure of explicit nonstationarity in the time-varying network (TVN) model, with certain unknown parameters, which are estimated iteratively using the delta-bar-delta training rule.[8] Provided that we can select the proper form of nonstationarity in the TVN model for a given application, this approach can provide a practical solution to a very important and difficult problem.

This paper presents the application of this general methodology to an actual neurophysiological system and demonstrates its efficacy in a practical context. The chosen nonstationary system is the dentate gyrus of the rabbit hippocampus under conditions of short-term potentiation.[2]

METHODOLOGY

The Volterra approach has been used for the modeling of nonlinear physiological systems, when no prior information is available about the internal characteristics of such systems or when the complexity of the system prevents the postulation of explicit parametric models.[12]

Address correspondence to V. Z. Marmarelis, OHE 500, USC, Los Angeles, CA 90089-1451. Electronic mail: vzm@bmsrs.usc.edu

It has been extended to the modeling of discrete-time nonlinear nonstationary systems,[12,13] whereby the input–output relationship is given by

$$y(n)=k_0(n)+T\sum_{m_1=0}^{M}k_1(n;m_1)x(n-m_1)+T^2\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}k_2(n;m_1,m_2)x(n-m_1)\times x(n-m_2)+\cdots, \quad (1)$$

where $x(n)$ denotes the input data sequence, $y(n)$ is the output data sequence, M is the finite memory of the system, T is the sampling interval, and $\{k_i\}$ are the Volterra kernels of the system that describe its nonlinear dynamics and are now time varying.

If we consider an expansion of the kernels on a complete temporal basis $\{f_j(n)\}$ over the data record $[0,R]$, then the Volterra series of Eq. (1) becomes

$$y(n)=\sum_j f_j(n)\left[\alpha_j^{(0)}+\sum_{m_1=0}^{M}\alpha_j^{(1)}(m_1)x(n-m_1)+\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\alpha_j^{(2)}(m_1,m_2)x(n-m_1)\times x(n-m_2)+\cdots\right], \quad (2)$$

where

$$k_1(n;m_1)=\sum_j f_j(n)\alpha_j^{(1)}(m_1), \quad (3)$$

$$k_2(n;m_1,m_2)=\sum_j f_j(n)\alpha_j^{(2)}(m_1,m_2), \quad (4)$$

etc. Equation (2) can be written as

$$y(n)=\sum_j f_j(n)y_j(n), \quad (5)$$

where

$$y_j(n)=\alpha_j^{(0)}+\sum_{m_1=0}^{M}\alpha_j^{(1)}(m_1)x(n-m_1)+\sum_{m_1=0}^{M}\sum_{m_2=0}^{M}\alpha_j^{(2)}(m_1,m_2)\times x(n-m_1)x(n-m_2)+\cdots. \quad (6)$$

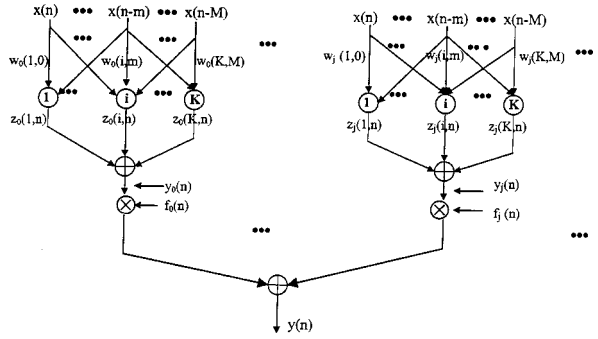


FIGURE 1. Time-varying network (TVN) architecture compatible with the general time-varying Volterra model. It is comprised of parallel feedforward subnets, with a single hidden layer having polynomial activation functions and an output modulating function (modulator). The modulators $\{f_j(n)\}$ represent the nonstationarity characteristics of the system. The stationary component of the system is represented by the subnet corresponding to $f_0(n)=1$.

Note that the signal $y_j(n)$ is the output of a time-invariant Volterra system with kernels $\{\alpha_j^{(i)}\}$ and input $x(n)$.

The basic mathematical relationships between time-invariant Volterra models and feedforward ANNs have been addressed by Marmarelis and Zhao,[15] and it has been shown that the two models become equivalent if the activation functions of the single hidden layer are allowed to be distinct polynomials with trainable coefficients. The specific relations between the Volterra kernels and the network parameters are given in Eqs. (A5)–(A7) of the Appendix. Note that other activation functions (of nonpolynomial form) can be selected as more physiologically meaningful in certain applications, as long as they form a complete basis. The equivalence with the Volterra model is defined through Taylor series expansions (if analytic) or Weierstrass polynomial approximations (if nonanalytic).

The proposed method employs the novel network architecture of Fig. 1 (termed TVN) to obtain models of time-varying Volterra systems using the delta-bar-delta training rule[6,8] for the training of the unknown network parameters. Each branch of the TVN in Fig. 1 produces an output that corresponds to each one of the terms in the summation of Eq. (5). Thus, the network configuration shown in Fig. 1 is equivalent to the time-varying Volterra model of nonstationary nonlinear systems, where the hidden units have distinct polynomial activation functions and the modulating functions $\{f_j(n)\}$ have characteristic parameters trainable with input–output data. We will refer to each branch of the TVN as a subnet. The novelty of the method is the introduction of a time-varying modulating function at the output of each subnet. A description of the TVN and the training rela-

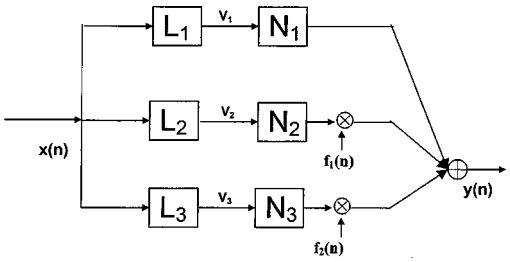


FIGURE 2. Nonlinear nonstationary system with one stationary and two nonstationary components used as a simulation example.

tions for the novel modulating functions are given in the Appendix.

The efficacy of this approach in modeling nonstationary, nonlinear systems has been demonstrated through computer simulations for several types of modulating functions (simply termed ''modulators''), such as transient, polynomial, periodic functions and combinations thereof.[7] In this paper, we present as an illustrative example the transient case of two sigmoidal modulators because it relates to the specific nonstationarity of the analyzed real data.

SIMULATION EXAMPLE

As an illustrative simulation example, we consider the time-varying second-order Volterra system of Fig. 2, where $L1$, $L2$, and $L3$ are linear filters and $N1$, $N2$, and $N3$ are static second-degree polynomial nonlinearities. The stationary path has the linear filter $L1$ and the static nonlinearity $N1$. The two nonstationary paths have two different linear filters: $L2$ and $L3$ followed by the static nonlinearities $N2$ and $N3$, respectively.

For the generation of the system data, we use 1024 point Gaussian white noise input. The first nonstationary path $(L2-N2)$ has the sigmoidal modulator $f_1(n) = 1/\{1+\exp[-0.03(n-250)]\}$ and the second one $(L3-N3)$ has the sigmoidal modulator $f_2(n)=1/\{1+\exp[-0.015(n-580)]\}$. The chosen linear filters have impulse response functions

$$L1: h_1(n) = 0.795039 \exp(-n/2),$$

$$L2: h_2(n) = 1.151676 n^2 \exp(-n), \tag{7}$$

$$L3: h_3(n) = 0.41889 n^2 \exp(-2n/3).$$

All the static nonlinearities are quadratic polynomials with unity coefficients and zero constant terms.

The corresponding TVN architecture for the training consists of one stationary subnet $[f_0(n)=1]$ and two nonstationary subnets with modulators $f_1(n)=1/\{1+\exp[-b_1(n-q_1)]\}$ and $f_2(n)=1/\{1+\exp[-b_2(n-q_2)]\}$, respectively. The training utilized the delta-bar-delta rule and yielded excellent estimates for the impulse responses $h1$, $h2$, and $h3$ after 200 iterations. The estimated parameters for the sigmoidal modulators and for the polynomial activation function are given in Table 1. Figure 3 shows the estimates for the impulse responses of the linear filters $L1$ (stationary path), $L2$ (first nonstationary path), and $L3$ (second nonstationary path), as well as the convergence patterns for the exponents b_1, b_2, the inflection points q_1, q_2, two representative weights, two representative polynomial coefficients, and the output root-mean-square (rms) error. It is evident that satisfactory convergence is achieved in 200 iterations.

REAL DATA ANALYSIS

The main goal of this paper is the application of the method to actual experimental data. The method was applied to neurobiological data consisting of extracellular population field potentials recorded from the dendritic layer of the dentate gyrus of the rabbit hippocampus. The experimental stimulus was a Poisson random sequence of impulses applied to the perforant path, the major excitatory input to granule cells, the principal neurons of the dentate gyrus. The Poisson random process had a mean rate of 5 impulses/s, which was sufficient to induce a progressive potentiation of perforant path-evoked granule cell response.[2,3,18] The data were sampled with a sampling frequency of 10 kHz, and the first 90 impulses were used for the analysis presented in this section. It was found that the nonstationary behavior was more evident in the first 90 pulses, and this fact was demonstrated by successfully predicting the response to the 30 pulses

TABLE 1. The exact and estimated parameters of the sigmoidal modulating functions and coefficients of the quadratic polynomial activation function of each hidden unit of the three subnets in the simulation example.

	Stationary subnet		First nonstationary subnet		Second nonstationary subnet	
	exact	estimate	exact	estimate	exact	estimate
Zeroth-order coefficient	0	0.136413*E*-06	0	0.523713*E*-05	0	0.948003*E*-05
First-order coefficient	1	0.9999739526	1	0.9999995326	1	1.0000180649
Second-order coefficient	1	0.9999455014	1	0.9999979664	1	1.0000292888
Exponents (b_1, b_2)	···	···	0.03	0.299999294	0.015	0.0149998286
Inflection points (q_1, q_2)	···	···	250	249.999799	580	579.999661

following the first 90 using the estimated model which exhibits essentially stationary behavior by the end of the 90 pulses. Because the power of the output signal was negligible above 3 kHz, this dataset was downsampled by a factor of 3 following lowpass filtering to avoid any aliasing.

Initially, we followed a piecewise stationary approach to explore whether this biological system exhibits a time-varying behavior. The input–output data were partitioned in five segments so that at least 800 samples (or almost 265 ms) elapse between the last impulse of each segment and the first impulse of the following segment in order to eliminate overlapping dynamics between segments. In addition to that, the minimum size of each segment was selected so that reasonable estimation accuracy can be achieved to follow changes in the dynamics of the system. The number of impulses in each segment were: 19 (10,500 sample points), 21 (9024 sample points), 20 (12,737 sample points), 14 (9700 sample points), and 16 (9380 sample points) in sequential order.

A stationary ANN with polynomial activation functions was used for the training. The selected parameters

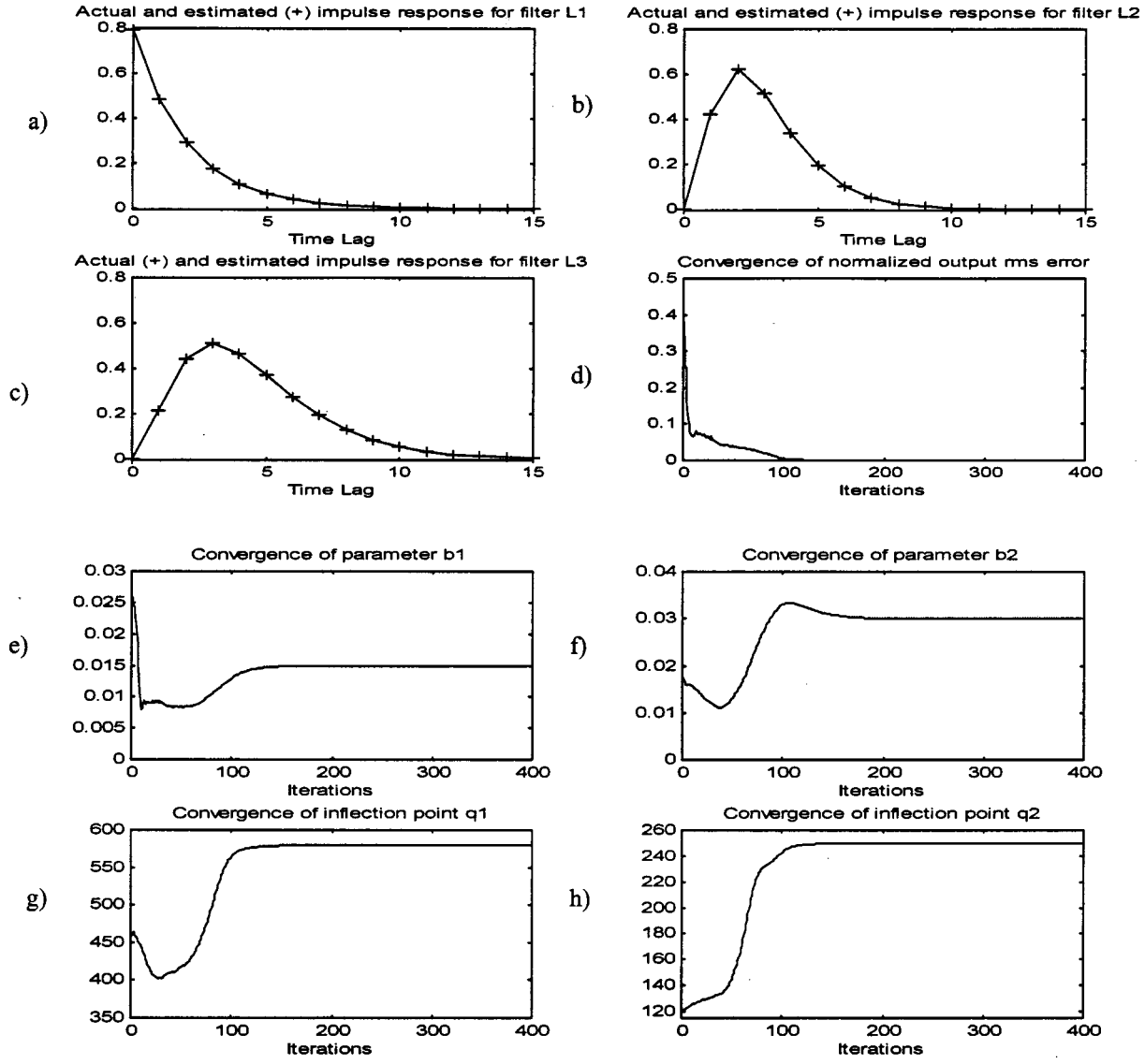


FIGURE 3. Results for the simulation example with one stationary and two nonstationary subnets with modulating functions: $f_1(n)=1/\{1+\exp[-b_1(n-q_1)]\}$ and $f_2(n)=1/\{1+\exp[-b_2(n-q_2)]\}$; (a), (b), (c) actual (solid line) and estimated (+) impulse responses for filters $L1$, $L2$, and $L3$, respectively; (d) the normalized output rms error vs training iterations; (e), (f) convergence patterns of the modulator exponents $b1$ and $b2$, respectively; (g), (h) convergence patterns of the modulator inflection points $q1$ and $q2$, respectively; convergence patterns of (i) the fifth weight of the stationary subnet, (j) the fifth weight of the first nonstationary subnet, (k) the second-degree coefficient of the stationary subnet and (1) the second-degree coefficient of the first nonstationary subnet. All parameters converge to the exact values within 200 iterations.

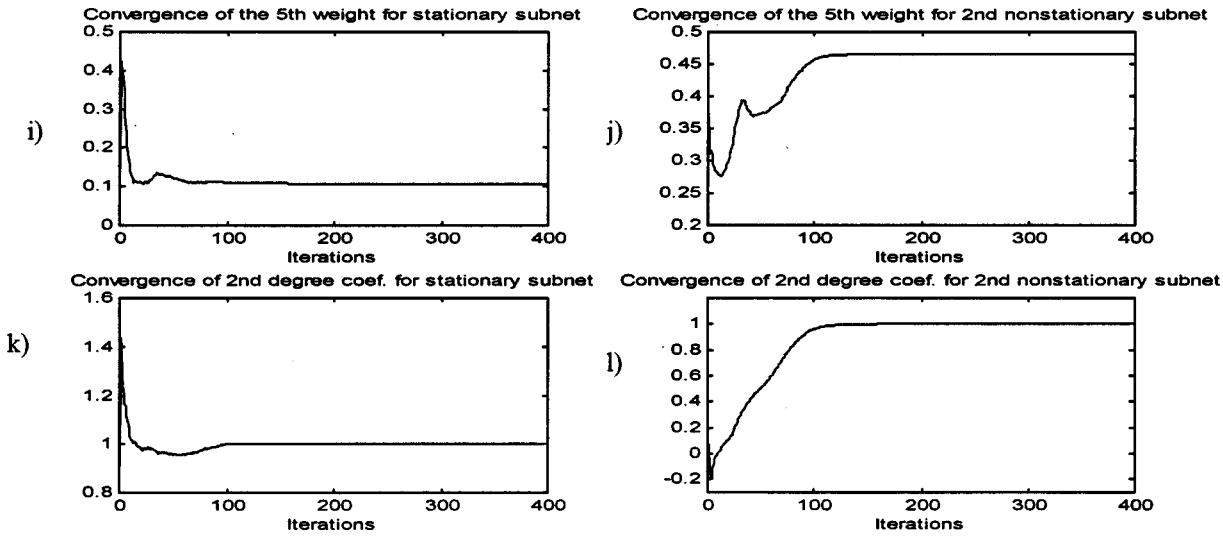


FIGURE 3. (*Continued.*)

were $M=400$, $K=3$, $R=2$ (i.e., 401 input units, three hidden units and second-degree polynomial activation functions). These parameters were chosen by successive trials using an output prediction error criterion. The total number of the estimated parameters is $[K\times(M+R+1)]=1209$, which is about 10%–14% of the number of datapoints in each segment, thus avoiding overfitting. The model predictions were good approximations of the real output signals for all segments as illustrated in Fig. 4, where representative segments are shown.

The responses to an impulse for each segment are presented in Fig. 5. Note that for a Volterra (nonlinear) system, the response to an impulse includes the diagonals of the high-order kernels in addition to the first-order (linear) kernel. An initial negativity occurs within the first five time lags for all segments, and represents the onset of an extracellular current sink (population excitatory postsynaptic potential) (EPSP) generated as a consequence of the intracellular depolarization of granule cell dendrites by perforant path input. The first positivity which follows represents an extracellular current source (compound action potential, or population spike) passively generated as a consequence of action potentials occurring in granule cell bodies located distally (thus, the reversed polarity). This positivity increases from the first to the third segment. A second positivity occurring within the first ten time lags, representing secondary action potential discharges, appears in the second segment and increases in amplitude from the second to the third segment. In the last two segments both positivities have almost equal amplitudes, which are smaller than the respective ones in the third segment. There is also a third peak within the same time period (first 16 ms), which appears in the third segment and may increase in amplitude in the next segment.

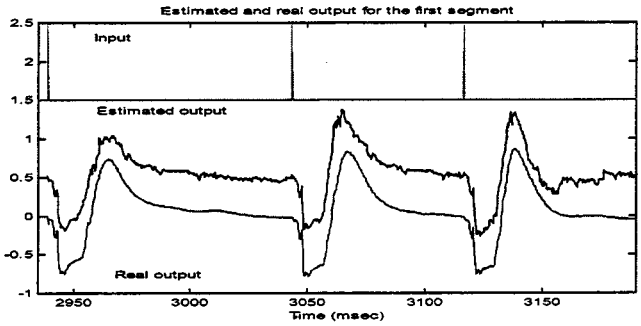
This preliminary analysis of the input/output data provided us with useful information regarding the way in which the system changed over time. The motivation was to incorporate this information in selecting the appropriate form of a global nonstationary model (i.e., the form of the ''modulators'' discussed previously). Again the parameters of the modulators would be determined from the training of the network.

The observation that the responses to an impulse exhibited distinct peaks that became more noticeable with time and remained almost unchanged during the last two segments led us to the selection of sigmoidal modulators, similar to those discussed in the Simulation Example section.

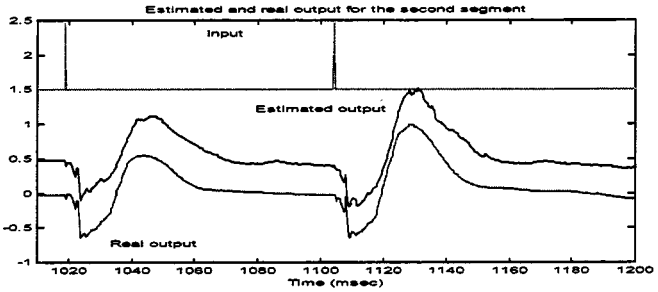
The fact that the first 16 ms of the responses to an impulse are of primary physiological interest, combined with our desire to facilitate the convergence of the algorithm by reducing the length, led us to the adoption of a surrogate set of input/output data. The surrogate output file consists of the first 50 sample points (almost 16.34 ms) of each of the five responses to an impulse (for each segment) and the input file is composed of the corresponding impulses at the beginning of each segment (Fig. 6). Notice that the abscissa is not real time, because of the elapsed time between segments. In real time, the first impulse is presented at 0 ms, the second at the 3,500 ms, the third at the 6,508 ms, the fourth at the 10,753 ms and the last one at the 13,987 ms. Each response to these input impulses lasts for 16.33 ms (or 50 sample points). Real time is used in the modulators of the TVN, but the training takes place only over the aforementioned limited surrogate data.

The TVN model used for training has one stationary subnet and two nonstationary subnets with the aforementioned sigmoidal modulators. All subnets have one hid-

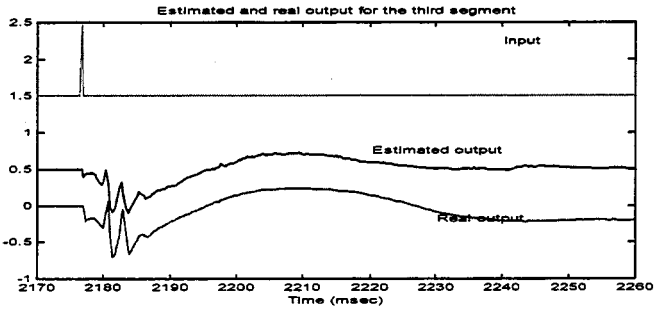
a)



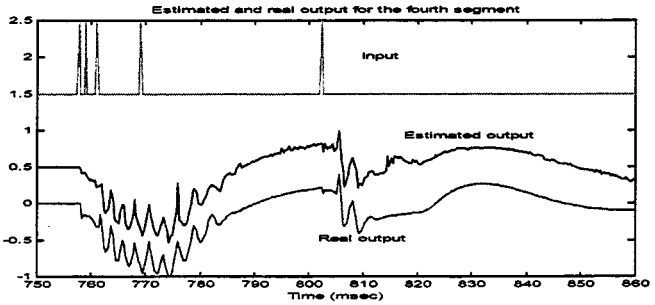
b)



c)



d)



e)

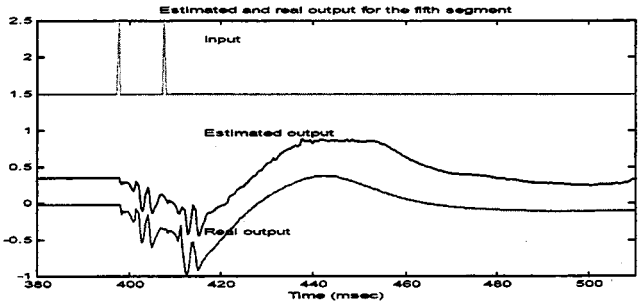


FIGURE 4. Illustrative portions of the experimental output for each of the five segments of the input/output data [(a)–(e)]. The predictive ability of the models is evident.

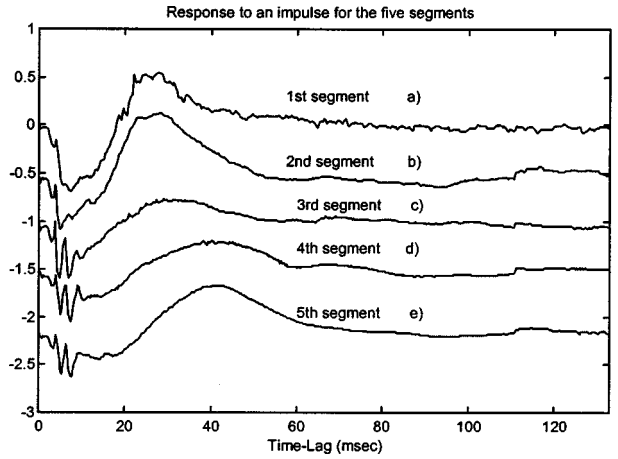


FIGURE 5. The model based response to an impulse for the first (a), second (b), third (c), fourth (d), and fifth (e) segments. The time-dependent changes due to potentiation are evident.

den unit, 50 input units and second-degree polynomial activation functions. The delta-bar-delta rule was used for training all parameters. The trained parameters were the 50 weights and the three coefficients of the activation functions (including the constants) for each subnet, and the parameters b_1, q_1, b_2, q_2 of the modulators.

The resulting estimates of the impulse responses of the corresponding linear filters of the stationary and nonstationary paths of the model are shown in Fig. 7. The estimated coefficients of the polynomial activation functions for each subnet and the parameters of the modulators are given in Table 2.

The output normalized rms error was 4.819%. The resulting modulators are depicted in Fig. 8. The upward transition of the first modulator begins at the beginning of the third segment and lasts for almost 5000 sample points (or 1.5 s). The downward transition of the second modulator begins at the end of the second segment and lasts for almost 2500 sample points (or 750 ms) entering the third segment.

Thus, the first nonstationary path is ''off'' during the first two segments and ''on'' during the last three segments (the transitional area extends over the first half of the third segment). The second nonstationary path is ''on'' during the first two segments and ''off'' during the last three segments (the transitional area extends briefly over the end of the second segment and the beginning of the third segment). The estimated output is depicted in Fig. 9. The estimated output for the last three segments is almost identical to the corresponding real output and constitutes a good prediction for the first two segments.

This result demonstrates the ability of the proposed methodology to model an important class of nonstationary systems with ''on'' and ''off'' switches that capture different states that the system may assume over time. Each switch will be a different subnet with a sigmoidal modulator introduced in the structure of the TVN model.

DISCUSSION

This strategy for identifying multiple nonstationarities with different characteristics raises the possibility of a mapping between the parallel paths of the TVN model and specific neurobiological mechanisms underlying the changes in the system state. The accuracy with which models achieved using the current implementation of the method represent both dominant and the subtler, detailed

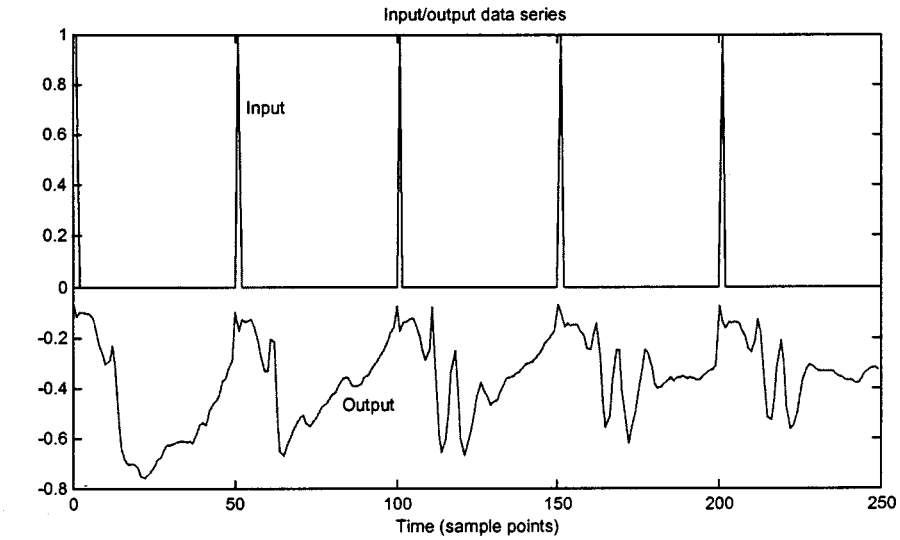


FIGURE 6. Surrogate input/output data for the training of the TVN model, which consists of one stationary and two nonstationary subnets with sigmoidal modulators. The output surrogate data is comprised of the first 50 sample points of each of the five responses to an impulse. The input surrogate data is comprised of the respective impulses.

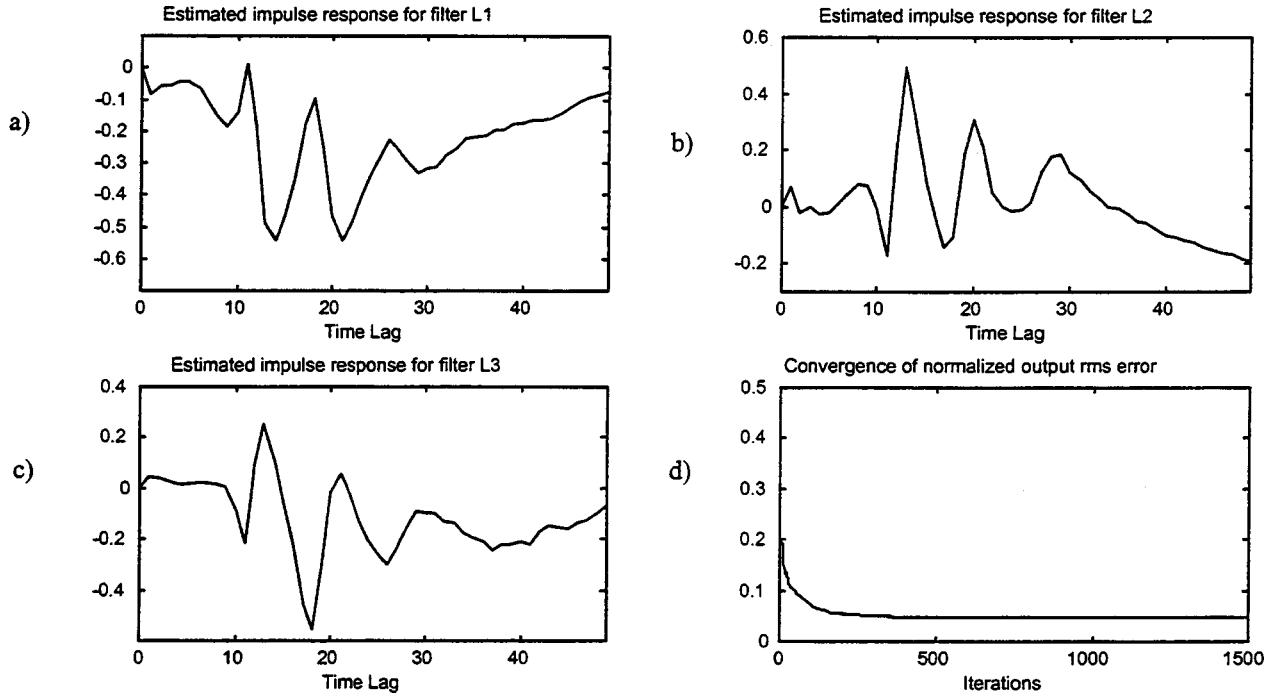


FIGURE 7. Results of the training of the TVN model with one stationary and two nonstationary subnets with modulating functions: $f_1(n)=1/\{1+\exp[-b_1(n-q_1)]\}$ and $f_2(n)=1/\{1+\exp[-b_2(n-q_2)]\}$. (a), (b), (c) Estimated impulse responses for filters *L*1, *L*2, and *L*3, respectively; (d) the normalized output rms error over 1,500 iterations.

alterations in the field potential wave form strengthens this possibility. To use the data analyzed in the present study as an example, the mean rate and intensity of stimulation used to activate perforant path fibers led to overlapping EPSPs and thus a progressive membrane depolarization of granule cells. The rate of depolarization is countered by input from inhibitory interneurons, but the prolonged depolarization leads to a failure of inhibitory circuits. Although admittedly a simplification, the second nonstationary path must reflect the collective dynamics of the temporal summation of granule cell EPSPs, voltage-dependent conductances activated by depolarization and which in turn amplify the depolarization further, and the progressive decline of inhibitory input. When inhibition ultimately fails, there is a rapid transition to a heightened level of depolarization, which activates additional, high-threshold excitatory conductances (e.g., voltage-dependent calcium channels) underlying the multiple population spikes. Stability is reached relatively rapidly, but the system emerges with decidedly different dynamics, the properties of which are captured by the first nonstationary path. The parameters for the modulators, b and q, should reflect the dynamics of the cellular processes underlying the transition from the low excitability state to the high excitability state, which in this case would be determined primarily by voltage-dependent conductances. Effects dependent on voltage-gated channels will evolve relatively rapidly (a time scale of seconds) compared to state transitions mediated by second messenger biochemical pathways that can have decay time constants of minutes or longer. For this reason, the steepness of the sigmoid functions shown in Fig. 8 is reasonable.

TABLE 2. The estimated parameters of the sigmoidal modulating functions and coefficients of the quadratic polynomial activation function of each hidden unit of the three subnets for the experimental data.

	Stationary subnet estimate	First nonstationary subnet estimate	Second nonstationary subnet estimate
Zeroth-order coefficient	−2.889714*E*-01	2.190199*E*-02	−1.411698*E*-02
First-order coefficient	1.137932	−8.164694*E*-01	−7.851913*E*-01
Second-order coefficient	2.706528*E*-01	3.782539*E*-02	5.077590*E*-02
Exponents (b_1, b_2)		1.684263*E*-03	2.346762*E*-03
Inflection points (q_1, q_2)		23,267.8(6.98 s)	18,790.8(5.64 s)

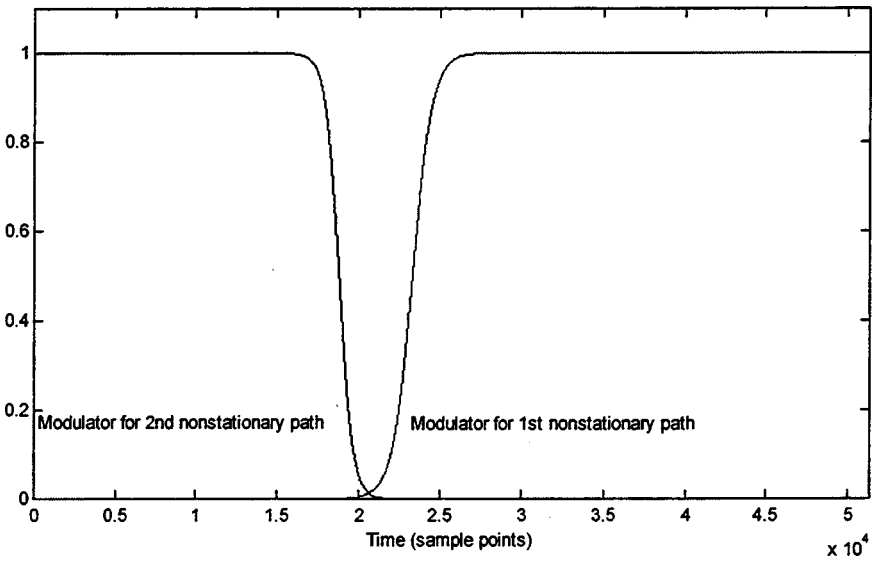


FIGURE 8. The estimated sigmoidal modulators for the two nonstationary subnets.

A state transition governed predominantly by much slower molecular and cellular processes should be associated with a smaller rate of change (assuming a sigmoid modulator remains appropriate) that would be less ''switch-like.'' In this manner, the dynamics of the two nonstationary paths and the properties of their modulators should place useful constraints on candidate mechanisms. A rigorous testing of this interpretational approach requires a substantially simplified neurobiological preparation having fewer available mechanisms. In such a simplified preparation, even subtle changes in the amplitude–time course of a bioelectric event are likely to be interpretable in terms of underlying molecular and cellular processes. This constitutes the next step in exploring the efficacy of this approach. Nonetheless, the presented results demonstrate the ability of the proposed methodology to extract useful information about the nonstationary dynamics of the system that is not currently accessible by any other methodological means.

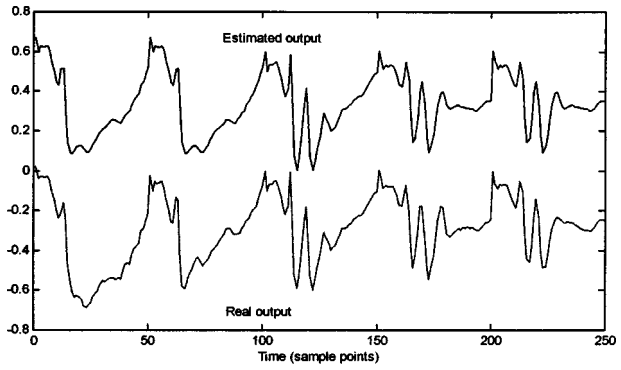


FIGURE 9. Experimental output (bottom trace) and prediction (top trace) for the model with one stationary and two nonstationary subnets.

ACKNOWLEDGMENTS

This work was supported by Grant No. RR-01861 awarded to the Biomedical Simulations Resource at the University of Southern California from the National Center for Research Resources and by Grant Nos. MH-51722 and MH-00343 from the National Institute of Mental Health of the National Institutes of Health.

APPENDIX

Each subnet has $(M+1)$ input units, receiving values corresponding to the stimulus epoch $[x(n)\ldots x(n-m)\ldots x(n-M)]$ at each time instant. Each hidden unit forms the weighted sum of the input values and transforms it nonlinearly by means of the respective activation function. For example, if we call $w_j(i,m)$ the weights for the jth subnet connecting the m-lag input value with the ith hidden unit, then the weighted sum

$$v_j(i,n)=\sum_{m=0}^{M} w_j(i,m)x(n-m) \tag{A1}$$

is formed and transformed nonlinearly by the respective polynomial activation function as

$$z_j(i,n)=\sum_{r=0}^{R} c_j(i,r)[v_j(i,n)]^r, \tag{A2}$$

where $z_j(i,n)$ is the output of the ith hidden unit of the jth subnet at time n. In Eq. (A2) $c_j(i,r)$ are the coefficients of the polynomial activation functions in the ith hidden unit of the jth subnet. Finally, assuming K hidden units for each subnet

$$y_j(n)=\sum_{i=1}^{K} z_j(i,n), \tag{A3}$$

$$y(n)=\sum_{j=0}^{L} f_j(n)y_j(n), \tag{A4}$$

where $y(n)$ is the model output, $y_j(n)$ is the output of the jth subnet and $f_j(n)$ is the modulating function of the jth subnet. For the stationary path $f_0(n)=1$.

Note that the kernels of each subnet in Eq. (6) can be expressed in terms of the weights and the polynomial coefficients of this subnet. For instance

$$a_j^{(0)}=\sum_{i=1}^{K} c_j(i,0), \tag{A5}$$

$$a_j^{(1)}(m)=\sum_{i=1}^{K} c_j(i,1)w_j(i,m), \tag{A6}$$

$$\alpha_j^{(2)}(m_1,m_2)=\sum_{i=1}^{K} c_j(i,2)w_j(i,m_1)w_j(i,m_2). \tag{A7}$$

If we assume that the modulator of the jth subnet has one unknown parameter p_j, then the following expression is used for updating this parameter:

$$\Delta p_j(k)=\alpha_p\Delta p_j(k-1) + \rho_p(k)[\tilde{y}(n)-y(n)]y_j(n)\frac{df_j(n)}{dp_j(k)}, \tag{A8}$$

where $\tilde{y}(n)$ is the measured output of the system, k denotes the update index and ρ_p, α_p are the initial learning step and momentum for the parameter p of the modulator, respectively. Note that the parameter updates take place at each time point n and after a complete presentation of the entire training set, we continue with successive iterations until the reduction in error becomes negligible. Thus, the update index k is distinct from the time index n and the number of iterations.

It is critical to note that the step size ρ (for every parameter) changes with k according to the delta-bar-delta rule[12]

$$\Delta\rho(k)=\begin{cases} k & \text{if } S(k-1)D(k)>0 \\ -\beta\rho(k) & \text{if } S(k-1)D(k)<0, \\ 0 & \text{otherwise} \end{cases} \tag{A9}$$

where $D(k)$ is the partial derivative of the error with respect to the parameter which is updated at time k, and $S(k)$ is an exponentially weighted sum of the current and past derivatives given by the first-order difference equation

$$S(k)=(1-\xi)D(k-1)+\xi S(k-1). \tag{A10}$$

Therefore, the step sizes evolve differently for the various parameters.

Using the chain rule of differentiation, we derive the following updating expressions for the polynomial coefficients of the activation functions:

$$\Delta c_j^{(k)}(i,r)=\alpha_c\Delta c_j^{(k-1)}(i,r) + \rho_c(k)[\tilde{y}(n)-y(n)]f_j(n)v_j^{(r)}(i,n) \tag{A11}$$

and for the in-bound weights of the hidden units:

$$\Delta w_j^{(k)}(i,m)=\alpha_w\Delta w_j^{(k-1)}(i,m) + \rho_w(k)[\tilde{y}(n)-y(k)]f_j(n)x(n-m) \times\left\{\sum_r rc_j(i,r)v_j^{r-1}(i,n)\right\}, \tag{A12}$$

where ρ_c, α_c and ρ_w, α_w are the initial learning step and momentum for the polynomial coefficients and the in-bound weights of each subnet of the TVN, respectively.

REFERENCES

[1] Arnold, M., W. H. R. Miltner, and H. Witte. Adaptive AR modeling of nonstationary time series by means of Kalman filtering. *IEEE Trans. Biomed. Eng.* 45:553–561, 1998.

[2] Berger, T. W., G. Barrionuevo, G. Chauvet, D. N. Krieger, and R. J. Sclabassi. A theoretical and experimental strategy for realizing a biologically based model of the hippocampus. In: Synaptic Plasticity. Molecular, Cellular, and Functional Aspects, edited by M. Baudry, R. F. Thompson, and J. L. Davis. Cambridge, MA: MIT Press, 1993, pp. 169–207.

[3] Berger, T. W., G. Barrionuevo, S. P. Levitan, D. N. Krieger, and R. J. Sclabassi. Nonlinear systems analysis of network properties of the hippocampal formation. In: Neurocomputation and Learning: Foundations of Adaptive Networks, edited by J. W. Moore and M. Gabriel. Cambridge, MA: MIT Press, 1991, pp. 283–352.

[4] Chen, S., S. A. Billings, and P. M. Grant. Nonlinear system identification using neural networks. *Int. J. Control* 6:1191–1214, 1990.

[5] Hassoun, M. H. Fundamentals of Artificial Neural Networks. Cambridge, MA: MIT Press, 1995, pp. 1–511.

[6] Haykin, S. Neural Networks. A Comprehensive Foundation. New York, NY: Macmillan, Press, 1994, pp. 1–696.

[7] Iatrou, M. Modeling of Nonlinear, Nonstationary Dynamic Systems with a Novel Class of Artificial Neural Networks.

Los Angeles, CA: University of Southern California, PhD dissertation, 1998.
[8] Jacobs, R. A., Increased rates of convergence through learning rate adaptation. *Neural Networks* 1:295–307, 1988.
[9] Krieger, D., T. W. Berger, and R. J. Sclabassi. Instantaneous characterization of time-varying nonlinear systems. *IEEE Trans. Biomed. Eng.* 39:420–424, 1992.
[10] Leontaritis, J., and S. A. Billings. Input–output parametric models for nonlinear systems. Part I: Deterministic nonlinear systems; Part II: Stochastic nonlinear systems. *Int. J. Control* 41:303–344, 1985.
[11] Marmarelis, P. Z., and V. Z. Marmarelis. Analysis of Physiological Systems: The White-noise Approach. New York, NY: Plenum, 1978, pp. 1–487.
[12] Marmarelis, V. Z. Practicable identification of nonstationary nonlinear systems. *IEE Proc. Part D,* , 128:211–214, 1981.
[13] Marmarelis, V. Z. Recent advances in nonlinear and nonstationary analysis. In: Advanced Methods of Physiological System Modeling, edited by V. Z. Marmarelis. New York, NY: Plenum, 1987, Vol. 1, pp. 323–336.
[14] Marmarelis, V. Z., Modeling methodology for nonlinear physiological systems. *Ann. Biomed. Eng.* 25:239–251, 1997.
[15] Marmarelis, V. Z., and X. Zhao. Volterra models and three-layer perceptrons. *IEEE Trans. Neural Netw.* 8:1421–1433, 1997.
[16] Narendra, K. S., and S. Mukhopadhyay. Adaptive control using neural networks and approximate models. *IEEE Trans. Neural Netw.* 8:475–485, 1997.
[17] Noriega, J. R., and H. Wang. A direct adaptive neural-network control for unknown nonlinear systems and its application. *IEEE Trans. Neural Netw.* 9:27–33, 1998.
[18] Thiels, E., G. Barrionuevo, and T. W. Berger. Excitatory stimulation during postsynaptic inhibition induces long-term depression in hippocampus *in vivo*. *J. Neurophysiol.* 72:3009–3016, 1994.
[19] Wray, J., and G. G. R. Green. Calculation of the Volterra kernels of nonlinear dynamic systems using an artificial neural network. *Biol. Cybern.* 71:187–195, 1994.