

EEG OCULAR ARTIFACT REMOVAL THROUGH ARMAX MODEL SYSTEM IDENTIFICATION USING EXTENDED LEAST SQUARES*

SHANE M. HAAS[†], MARK G. FREI[‡], IVAN OSORIO[§], BOZENNA PASIK-DUNCAN[¶],
AND JEFF RADEL^{||}

Abstract. The removal of ocular artifact from scalp electroencephalograms (EEGs) is of considerable importance for both the automated and visual analysis of underlying brainwave activity. Traditionally, subtraction techniques use linear regression to estimate the influence of eye movements on the electrodes of interest. These methods are based on the assumption that the underlying brainwave activity is uncorrelated when, in general, it is not. Furthermore, regression methods assume that the ocular artifact propagation is frequency independent, i.e. all waveforms of the ocular artifact propagate similarly. In this paper, we examine relaxing these assumptions by using a more general autoregressive (AR) moving average (MA) exogenous (X) model and the extended least squares (ELS) algorithm to remove ocular artifact. We demonstrate that in some cases this general ARMAX model can decrease ocular artifact not removable by standard regression techniques. We also show that the incorporation of a forgetting factor to exponentially weight past data can improve ocular artifact removal even for the traditional subtraction method.

Keywords. Ocular artifact removal; System identification; Extended least squares; Adaptive noise cancellation; Epilepsy; Electroencephalogram signal processing; Weighted least squares

1. Introduction.

1.1. Ocular Artifact in EEGs. Ocular artifact (OA) refers to the electromagnetic potentials superimposed on EEGs associated with eye movements (EM) and/or blinks. These potentials originate from the electric dipole between the cornea and retina and depend on its orientation with respect to the eyelid [1], [2]. These larger artifacts sometimes obscure the smaller amplitude underlying brainwave activity, making its visual or automated inspection difficult. The goal of ocular artifact removal (OAR) methods is to remove ocular artifacts from scalp recorded EEGs, leaving the underlying background signals due to brain activity.

*Received on April 15, 2002; accepted for publication on November 22, 2002.

[†]Laboratory for Information and Decision Systems, and Research Laboratory of Electronics, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge MA 02139. E-mail: shaas@alum.mit.edu

[‡]Corresponding Author. Flint Hills Scientific, L.L.C. 5020 W. 15th St., Suite A Lawrence, KS 66049. E-mail: frei@sound.net

[§]Comprehensive Epilepsy Center, 4015 Bell Kansas University Medical Center 3901 Rainbow Blvd. Kansas City, KS 66160-7314 and Flint Hills Scientific, L.L.C. 5020 W. 15th St., Suite A Lawrence, KS 66049. E-mail: iosorio@kumc.edu

[¶]Dept. of Mathematics, 405 Snow Hall University of Kansas Lawrence, KS 66045. E-mail: bozenna@math.ukans.edu

^{||}Mail Stop 3051 (1018 MRRC) Univ. Kansas Medical Center 3901 Rainbow Blvd. Kansas City, KS 66160-7336. E-mail: jradel@kumc.edu

Prior to the introduction of automated OAR methods, subjects were asked to stare at a fixed object or picture to reduce eye movements. Another early solution was simply to discard data corrupted with severe artifact. Neither of these methods produced satisfactory results. The latter produced large amounts of lost data [3], and the former did not account for blinks or small involuntary eye movements [4]. Automated OAR methods, like the subtraction method discussed in the next section, allow for online removal of ocular artifacts. For a survey of OAR methods see [5].

Figure 1 shows a segment of EEG signals corrupted with ocular artifact. Notice the large dips on channels FP1-F3, FP2-F4, FP1-FP7, and FP2-F8. The blink artifacts are so prominent on these channels because they are located nearest to the eyes (Figure 2(a)). Subtraction OAR methods use electrodes placed near the eyes to measure eye movements and then subtract them from the EEG.

We will examine and expand on these subtraction OAR methods in this paper, which is structured as follows. The remaining introductory sections explain the subtraction OAR technique, and then extend this method to a more general ARMAX model. The extended least squares and order estimation algorithms are also introduced to estimate the parameters of the ARMAX model. Section 2 describes our data collection procedures and techniques to evaluate the performance of OAR. Section 3 demonstrates our ARMAX model subtraction OAR method and compares them to the standard linear regression subtraction method.

1.2. Subtraction Ocular Artifact Removal Methods. Subtraction methods are based on the assumption that the measured EEG is described as a linear combination of the underlying cortical activity, or background EEG, and the corrupting ocular artifact [5]. They generally use a variation of the following model to relate the measured EEG, background EEG, and ocular artifact

$$(1) \quad y(n) = \sum_{i=1}^N b_i u_i(n) + w(n),$$

where $y(n) \in \mathcal{R} = \{ \text{Set of Real Numbers} \}$ is a single channel/electrode of measured EEG at time n , $u_i(n) \in \mathcal{R}, i \in \{1, \dots, N\}$ is the ocular artifact as measured from electrodes placed near the eyes called electrooculograms (EOG), and $w(n) \in \mathcal{R}$ is the true background EEG. The integer N varies from method to method, depending on how the signals u_i are defined. The constants $b_i \in \mathcal{R}, i \in \{1, \dots, N\}$ are sometimes called transmission coefficients, and they determine how much OA is prevalent in the measured EEG $y(n)$. Subtraction OAR methods estimate these coefficients, and then subtract the ocular artifact from the measured EEG to obtain the background EEG. The background EEG, which can be viewed as an error term, is then assumed to be the underlying brain activity. These subtraction methods can also be viewed in the framework of the adaptive noise cancellation problem, and are discussed as such in [6] and [7].

Subtraction OAR methods vary in how the ocular artifact signals, $u_i(n), i \in \{1, \dots, N\}$, are defined. These definitions are sometimes chosen to measure certain types of eye movements such as horizontal and vertical motions or nonlinear interactions between ocular dipoles. Another consideration for choosing them is to minimize the number of electrodes required for OAR. Figure 2(b) shows frequently used electrode placements for measuring eye movements. The voltage signals are measured with respect to a reference generated by electrodes placed at the mastoids located behind each ear. Instead of nontrivial linear combinations and/or products of signals obtained from the electrodes near the eyes we instead use the raw signals themselves for the exogenous inputs. Doing so encompasses most previous models found in the literature (see, e.g., [5], [8], [9], [10], [11]).

The regression model of (1) usually assumes the following:

- A1** The background EEG $w(n)$ is uncorrelated (white) noise with zero mean ([5], [11]);
- A2** All frequencies of the EOG have the same propagation characteristics. In other words, $y(n)$ is not dependent on $u_i(t)$ for $t < n$ ([12], [11], [5]);
- A3** The EOG $u_i(n), i \in \{1, \dots, N\}$ and background EEG $w(n)$ are uncorrelated ([12], [5]);
- A4** All types of eye movements (e.g., blinks, saccades, etc.) have the same propagation characteristics (i.e., transmission coefficients) ([5]);
- A5** The transmission coefficients do not vary with time ([5]).

For least squares estimation of the transmission coefficients, these assumptions are necessary for the estimates to converge in probability to the true coefficients [13]. Without **A1**, the estimates might no longer converge in probability, but still remain equal to the true values in expectation [14]. In practice **A1** does not hold because the underlying brainwave activity is significantly correlated [15], [16], [17]. To account for this correlation, van den Berg-Lenssen in [11] proposed to model the background EEG as the moving average of uncorrelated noise, i.e. (1) becomes

$$(2) \quad y(n) = \sum_{k=1}^N b_k u_k(n) + \sum_{k=1}^r c_k w(n-k) + w(n),$$

where $w(n)$ is assumed to be uncorrelated and mean zero. Jarvis in [18] recommended using differenced data to overcome background EEG correlation. Differenced data are obtained by subtracting the value of the data at time $n-1$ from that at time n . Using differenced data leads to the model

$$(3) \quad y(n) = y(n-1) + \sum_{k=1}^N b_k (u_k(n) - u_k(n-1)) + w(n) - w(n-1),$$

which is a special case of the ARMAX model discussed in the next section. Tuan noted in [19], however, that using differenced data may reduce autocorrelation, but only removes it when the process is first order autoregressive with a unit coefficient.

To relax assumption **A2**, Jervis in [5], van den Berg-Lenssen in [11], and Sadasivan in [6] incorporated past EOG measurements in modeling the current measured EEG, i.e. (1) becomes

$$(4) \quad y(n) = \sum_{k=0}^q \sum_{i=1}^N b_{ki} u_i(n-k) + w(n),$$

where $b_{ki}, i \in \{1, \dots, N\}, k \in \{1, \dots, q\}$ are the transmission coefficients and q is a positive integer [19]. This model is also a special case of the more general ARMAX model presented in the next section.

In practice, assumption **A3** does not hold because the EOG contains a small amount of background EEG and other non-ocular artifacts. Background EEG is most prevalent on the EOG electrodes placed above the eyes; therefore, Jervis in [18] recommends using the smallest number of these electrodes for OAR subtraction. Using too few, however, might result in bias due to omitting necessary regressors [20].

Also, all types of eye movements do not propagate to the scalp electrodes in the same manner due to their physiological origin, violating assumption **A4**. For example, ocular artifact from small deflection eye movements originates primarily from deflections of the two ocular dipoles. Ocular artifact, however, from voluntary blinks originate primarily from interactions between the dipoles, eyelids, and muscles surrounding the eye. Gratton *et al.* noted this observation in [21] showing the differences between transmission coefficients estimated from voluntary and involuntary eye movements. To distinguish between different types of eye movements, and build appropriate models for each, is a very time-consuming procedure, and usually not suitable for online applications [5]. For this reason, we will generally assume that **A4** holds, but take into account the possible time-varying behavior by introducing a “forgetting factor” weighting in our estimation algorithm (see, e.g., [22]).

Gratton also showed in [21] that for a given subject, the parameter estimates were not constant during or between experimental sessions. They also noted that the variation between sessions was larger than that during the same session. These results contradict assumption **A5**, and suggest that an OAR method needs to adapt over time to track the transmission coefficients. Introducing a forgetting factor [23] or weighting [24] to the recursive least squares algorithm can place more emphasis on recent data, allowing the estimator to track time-varying coefficients. Other algorithms such as least mean squares (LMS) [25] or stochastic approximation [26] can also track a time varying system.

1.3. ARMAX Models and Their Application to OAR. We will now generalize models (1), (2), (3), and (4) used in the previously mentioned studies to relax assumptions **A1** and **A2** by modeling the measured EEG as an ARMAX process

described as

$$(5) \quad y_n = \sum_{k=1}^p A_k y_{n-k} + \sum_{k=0}^q B_k u_{n-k} + \sum_{k=1}^r C_k w_{n-k} + w_n,$$

where we represent all measured EEG channels as a d_1 -dimensional vector $y_n \in \mathcal{R}^{d_1}$, all measured EOG channels as a vector $u_n \in \mathcal{R}^{d_2}$, and $w_n \in \mathcal{R}^{d_1}$ is a mean zero uncorrelated process generating the background EEG. For convenience, we will adopt the subscript notation to represent time. The model order is defined as the triplet (p, q, r) . An ARMA, ARX, and XMA model would then have orders expressed as $(p, -1, r)$, $(p, q, 0)$, and $(0, q, r)$, respectively. The coefficients of the model are matrices, $A_k \in \mathcal{R}^{d_1 \times d_1}$, $k \in 1, \dots, p$, $B_k \in \mathcal{R}^{d_1 \times d_2}$, $k \in 1, \dots, q$, and $C_k \in \mathcal{R}^{d_1 \times d_1}$, $k \in 1, \dots, r$.

Notice that the models presented in the previous section are special cases of (5). Models (1), (2), (3), and (4) are $(0, 0, 0)$, $(0, 0, r)$, $(1, 1, 1)$, and $(0, q, 0)$ ARMAX models, respectively. Thus, an ARMAX model encompasses many of the models proposed by other authors.

To extract the background EEG from (5), we must examine how w_n and u_n become stored in the autoregressive terms. Define

$$(6) \quad A(z) = A_1 + A_2 z + \dots + A_p z^{p-1},$$

$$(7) \quad B(z) = B_0 + B_1 z + \dots + B_q z^q,$$

$$(8) \quad C(z) = I + C_1 z + \dots + C_r z^r,$$

where z denotes the backshift operator. We can then write (5) as

$$(9) \quad y_n = A(z)y_{n-1} + B(z)u_n + C(z)w_n.$$

Assuming $y_n = 0, n < 0$ and iterating gives

$$(10) \quad y_n = A^n(z)B(z)u_0 + A^{n-1}(z)B(z)u_1 + \dots + B(z)u_n + v_n,$$

where we have defined v_n as

$$(11) \quad v_n = \begin{cases} 0 & n < 0 \\ A(z)v_{n-1} + C(z)w_n & n \geq 0 \end{cases},$$

or, equivalently, as

$$(12) \quad v_n = \begin{cases} 0 & n < 0 \\ A(z)v_{n-1} + y_n - A(z)y_{n-1} - B(z)u_n & n \geq 0 \end{cases}.$$

The ARMA process v_n is the portion of y_n without ocular corruption, i.e. the background EEG. From (12), we see that only $A(z)$ and $B(z)$ need to be estimated to calculate the background EEG. The next section shows how to estimate the matrix coefficients of (5) using the extended least squares algorithm.

1.4. ARMAX Model Identification Using Extended Least Squares. Define the true coefficients of (5) as $\theta \in \mathcal{R}^{(p+r)d_1+(q+1)d_2 \times d_1}$

$$(13) \quad \theta^T = [A_1 \cdots A_p \quad B_0 \cdots B_q \quad C_1 \cdots C_r],$$

and the true regressor $\psi_n \in \mathcal{R}^{(p+q)d_1+(q+1)d_2}$ as

$$(14) \quad \psi_n^T = [y_n^T \cdots y_{n-p+1}^T \quad u_{n+1}^T \cdots u_{n-q+1}^T \quad w_n^T \cdots w_{n-r+1}^T],$$

where T denotes the matrix transpose operator. Then (5) becomes

$$(15) \quad y_n = \theta^T \psi_{n-1} + w_n.$$

Define the estimated regressor vector $\phi_n \in \mathcal{R}^{(p+q)d_1+(q+1)d_2}$ as

$$(16) \quad \phi_n^T = [y_n^T \cdots y_{n-p+1}^T \quad u_{n+1}^T \cdots u_{n-q+1}^T \quad \hat{w}_n^T \cdots \hat{w}_{n-r+1}^T],$$

where

$$(17) \quad \hat{w}_n = y_n - \hat{\theta}_n^T \phi_{n-1},$$

and $\hat{\theta}_n$ is the estimate for θ to be defined shortly. Denote the collection of past data as

$$(18) \quad Y_n^T = [y_1 \cdots y_n],$$

$$(19) \quad \Phi_n^T = [\phi_0 \cdots \phi_n].$$

The extended least squares estimate at time n denoted as $\hat{\theta}_n$ is

$$(20) \quad \hat{\theta}_n = \arg \min_{\theta} \|(Y_n - \Phi_{n-1}\theta)^T R_n (Y_n - \Phi_{n-1}\theta)\|,$$

where $\hat{\theta}_0$ is chosen arbitrarily and R_n is a symmetric matrix of weighting coefficients. Based on modifications from results in [26] and [27] the extended least squares estimate is

$$(21) \quad \hat{\theta}_n = (\Phi_{n-1}^T R_n \Phi_{n-1})^{-1} \Phi_{n-1}^T R_n Y_n,$$

assuming that the inverse of $\Phi_{n-1}^T R_n \Phi_{n-1}$ exists. In practice, we will ensure this invertibility by adding a small positive multiple of the identity matrix to $\Phi_{n-1}^T R_n \Phi_{n-1}$. If $R_n = \text{diag}\{a_0 \lambda^n, \dots, a_{n-1} \lambda, a_n\}$, where $a_n \in \mathcal{R}$ is a weighting sequence and $0 < \lambda \leq 1$ is the forgetting factor, then we can write

$$(22) \quad \hat{\theta}_n = \left(\sum_{k=1}^n \lambda^{n-k} a_k \phi_{k-1} \phi_{k-1}^T + \frac{\lambda^n}{\alpha} I \right)^{-1} \sum_{k=1}^n \lambda^{n-k} a_k \phi_{k-1} y_k^T,$$

where $\alpha \in \mathcal{R}$, and is included to ensure invertibility. Define

$$(23) \quad P_n = \left(\sum_{k=1}^n \lambda^{n-k} a_k \phi_{k-1} \phi_{k-1}^T + \frac{\lambda^n}{\alpha} I \right)^{-1}.$$

Then, the recursive extended least squares estimator with exponential forgetting factor $0 < \lambda \leq 1$ and weighting sequence $a_n \in \mathcal{R}$ can be recursively written as

$$(24) \quad \hat{\theta}_n = \hat{\theta}_{n-1} + a_n P_n \phi_{n-1} \left(y_n^T - \phi_{n-1}^T \hat{\theta}_{n-1} \right), \text{ and}$$

$$(25) \quad P_n = \left(\lambda P_{n-1}^{-1} + a_n \phi_{n-1} \phi_{n-1}^T \right)^{-1}$$

$$(26) \quad = \lambda^{-1} P_{n-1} - \frac{\lambda^{-1} P_{n-1} \phi_{n-1} \phi_{n-1}^T P_{n-1}}{\lambda/a_n + \phi_{n-1}^T P_{n-1} \phi_{n-1}},$$

with P_0 chosen as αI . The derivation of this estimator is based on a combination of proofs given by [23], [26], [28], and [24].

We will now examine the conditions necessary for strong consistency of the extended least squares estimator. Define

$$(27) \quad r_n = \text{tr} \{ P_n^{-1} \} = \text{tr} \{ \alpha I \} + \sum_{k=1}^n \lambda^{n-k} a_k \phi_{k-1}^T \phi_{k-1}.$$

With $\lambda = 1 = a_n \forall n$, Chen and Guo in [29] showed strong consistency of the ELS estimate under the following assumptions:

A6 The noise $\{w_n, F_{n-1}\}$ is a martingale difference sequence with respect to a family of non-decreasing σ -algebras $\{F_n\}$ and $\sup_{n \geq 0} E \{ \|w_n\|^\beta | F_{n-1} \} < \infty$ a.s. (almost surely) for some $\beta \geq 2$;

A7 u_n is F_{n-1} -measurable;

A8 $C(z)^{-1} - \frac{1}{2}I$ is strictly positive real (SPR) where $C(z) = I + C_1 z + \dots + C_r z^r$.

Under the above assumptions, if $\beta > 2$ then

$$(28) \quad \|\hat{\theta}_n - \theta\| = o \left(\left(\frac{\log r_n}{\lambda_{min}^n} \right)^{\frac{1}{2}} \right) \text{ a.s.},$$

where λ_{min}^n is the minimum eigenvalue of P_n^{-1} . We use the notation $f_n = o(g_n)$ for two sequences f_n and g_n to mean that $\lim_{n \rightarrow \infty} f_n/g_n = 0$. If $\beta = 2$ then for any $c > 1$

$$(29) \quad \|\hat{\theta}_n - \theta\| = o \left(\left(\frac{\log r_n (\log \log r_n)^c}{\lambda_{min}^n} \right)^{\frac{1}{2}} \right) \text{ a.s.}$$

In practice, assumption **A6** means that w_n is mean zero, uncorrelated, and has a finite higher moment. Taken together, **A6** and **A7** imply that u_n and w_n are uncorrelated, i.e. assumption **A3**. We can interpret **A8** as saying the moving average noise is not too different from white noise [13]. Under similar assumptions, Bercu in [24] and Guo in [30] showed strong consistency of the weighted least squares estimator for $\lambda = 1$ and a_n satisfying certain conditions of admissibility. Solo in [23] illustrated the estimator's behavior for $a_n = 1 \forall n$ and $0 < \lambda \leq 1$.

For the purpose of OAR, we do not require the estimation of $C_k, k \in \{1, \dots, r\}$ because by (12) we can estimate the underlying background EEG with estimates of

$A_k, k \in \{1, \dots, p\}$ and $B_k, k \in \{1, \dots, q\}$. This observation is important because we are unable to verify assumption **A8** for EOG and EEG data and believe that it is unlikely to be satisfied for these data. Although we do not know the true model for this data, the characteristic roots of the estimated polynomial for $C(z)$ are inside the unit circle, which violates the SPR condition [31]. So we wish to relax assumption **A8**, yet still estimate the autoregressive and exogenous coefficients. We have conjectured and confirmed through simulations that under assumptions **A6** and **A7**, but not **A8**, that the autoregressive and exogenous coefficient estimates converge to their true values, but the moving average coefficient estimates converge with a bias. If our conjecture is true, as suggested by simulation, estimates of the background EEG using the ELS should be representative of the true background regardless of conditions on $C(z)$.

1.5. Order Estimation. The choice of ARMAX model order is a very difficult problem because the actual physiological mechanism for ocular artifact propagation is unknown. Hence, we are attempting to fit a finite order, causal, time-invariant, linear model to a possibly infinite order, non-causal, time-variant, non-linear phenomenon. The only justification for doing so, stems from the observation that the ocular artifacts in the measured EEG seem to be of similar shape and occur at similar times to those appearing in the measured EOG. Similar to approximating a continuous function locally with a finite order polynomial, we hope to approximate the actual mechanism with our finite order model.

If this mechanism really behaves as an ARMAX model, then offline methods do exist to consistently estimate its order. Chen and Guo in [26] offer order estimation techniques based on the CIC information criterion for ARMAX models with and without *a priori* bounds on the model order. Other information criterion are presented in [32], [33], and [34] to estimate the order of an ARMA model. All these methods, however, require searching over a large number of data points and model orders to minimize the information criterion. Furthermore, to show strong consistency of the order estimates requires complicated assumptions that are not verifiable for unknown systems.

These methods, however, provide a loose framework for us to choose the model order. We will use a modification of the CIC information criterion discussed in [26] and [14]. We will choose (p, q, r) to minimize

$$(30) \quad CIC(p, q, r)_n = \sigma_n(p, q, r) + (p + q + r + 1) \log(n - n_0)\gamma,$$

where

$$(31) \quad \sigma_n(p, q, r) = \sum_{k=n_0}^n \|\hat{w}_k\|^2,$$

$$(32) \quad \gamma = \text{var}\{w_n\},$$

\hat{w}_k is defined in (17), and n_0 is a delay or starting time. We also have explicitly included the dependency of the estimator (22) on the selected order (p, q, r) in the previous equations. The cumulative error in (31) is a measure of how well a (p, q, r) model “fits” the observed data. The second term in (30) penalizes higher order models based on the “principle of parsimony” of model-building, meaning that all other things being equal, the model with the smallest number of parameters is preferable [14]. The positive constant γ scales this second term so it is of comparable magnitude to the first. Its inclusion is necessary when dealing with a finite amount of data because $\sigma_n(p, q, r)$ might be less than $\log n$ for all orders, which would result in selection of the trivial $(0, -1, 0)$ model. When dealing with EEG data, we choose γ to be the variance of a channel not significantly corrupted with ocular artifact, such as the one located farthest from the eyes, because we have found this choice to effect a reasonable scaling. The starting time, n_0 , reduces the estimator’s transient effects on the cumulative error. When estimating the order for the EEG and EOG model, we chose n_0 to be approximately half the length of the data set.

We will use the search mechanism proposed in [26] to minimize $CIC(p, q, r)_n$. This mechanism does not search over all possible order triplets, but instead determines a bound on the order then proceeds to estimate each order in the triplet individually as follows:

$$(33) \quad \hat{m}_n = \arg \min_{0 \leq k \leq \log n} CIC(k, k, k)_n,$$

$$(34) \quad \hat{p}_n = \arg \min_{0 \leq p \leq \hat{m}_n} CIC(p, \hat{m}_n, \hat{m}_n)_n,$$

$$(35) \quad \hat{q}_n = \arg \min_{0 \leq q \leq \hat{m}_n} CIC(\hat{p}_n, q, \hat{m}_n)_n,$$

$$(36) \quad \hat{r}_n = \arg \min_{0 \leq r \leq \hat{m}_n} CIC(\hat{p}_n, \hat{q}_n, r)_n.$$

To verify that the order estimates are reasonable, in this search’s implementation we record the two terms in (30). Afterwards, we can interpret the estimator’s choice based on the minimization of each term separately.

We evaluated the performance of this order estimator through simulation. In general, this method seems to produce reasonable estimates of the model order, but are not always correct due to finite data sets and the inapplicability of assumptions required for consistent estimation.

Another approach to estimating model order is to over-parameterize the model. For example, if the true model order is, say $(4, 4, 4)$, then estimating the coefficients assuming a $(6, 6, 6)$ model would yield estimates with the two higher order coefficients approximately zero. Model order could then be chosen as the highest order with approximately non-zero coefficients. This method has many practical flaws. First, it assumes an upper bound on the model order. Second, determining which coefficients are negligible is not well defined. Third, as the estimator order increases, its variance does as well, resulting in less certain estimates. An advantage of this method, however,

is that it is quicker than performing an extended search to minimize an information criterion.

Another consideration when choosing model order is its physical interpretation. From model (10) with order (p, q, r) we see that the underlying brainwave activity is described by the ARMA process (11) of order $(p, -1, r)$. If the noise w_n of this ARMA process is not white, then a bias will result in the ARMAX model estimates [14]. A bias in the estimates will result in too much or too little subtraction of ocular artifact, resulting in poor removal. Hence, p and q must be chosen so that the underlying brainwave activity, or background EEG, is appropriately modeled as an ARMA process with white noise excitation.

2. OAR Methodology. This section discusses the methods we used to collect the EEG and EOG data and evaluate the performance of the OAR methods.

2.1. Data Collection. A BMSI System 4000 (Los Gatos, CA) collected the EEG and EOG data using a 239.75 Hz sampling rate with 10 bits of precision across a $\pm 300\mu V$ range. Prior to sampling, the signals were filtered by anti-aliasing filters with nominal -3 dB passbands of 0.3-70 Hz. After digitization, we low-pass filtered the data with a digital Butterworth filter having a 50 Hz -3 dB cutoff frequency with at least a -100 dB attenuation for frequencies greater than or equal to 60 Hz. This filter removed the 60 Hz noise that was very prevalent in the data. The data were collected from two subjects for approximately 30 minutes each. During this time, the subjects performed various natural and prompted eye movements such as blinks, saccades, fixation, and smooth pursuit movements. We then selected segments of data exhibiting large amounts of ocular artifact but little muscle artifact for input into the OAR algorithm.

2.2. Evaluating the Performance of OAR Methods. No standard exists to quantitatively evaluate the performance of OAR methods [8]. Several methods, however, can qualitatively assess OAR performance. The most obvious method is visual inspection. Ocular artifact has distinct waveshapes recognizable by a trained observer. After OAR, these waveshapes are either removed or diminished; therefore, visual inspection provides an initial measure of OAR performance. Jervis in [35] suggests, however, that visual comparison of EEG before and after OAR may not be sufficient to fully evaluate performance.

Another commonly used method to evaluate OAR performance is the minimization of the cumulative error in (31) [11]. After all, this is the error that the least squares method tries to minimize. For the subtraction OAR model of (1), minimization of (31) is equivalent to the minimization of the estimated background EEG variance. The weakness of this OAR evaluation procedure is its reliance on assumption **A3**, i.e., that EOG and EEG are uncorrelated. When this assumption fails, the subtraction OAR method not only removes ocular artifact from the measured EEG,

but background EEG as well. The unintentional removal of background EEG reduces the cumulative error (31), yet is undesirable.

Similar to the previous method, the statistical properties of \hat{w}_n can also indicate OAR performance [11]. If assumption **A1** is true, i.e., background EEG is white, and \hat{w}_n converges to w_n , then \hat{w}_n should be uncorrelated. If we examine its sample autocorrelation function,

$$(37) \quad R_{\hat{w}\hat{w}}(\tau) = \frac{1}{n} \sum_{k=1}^n \hat{w}_k \hat{w}_{k+\tau},$$

it should be approximately zero for $\tau \neq 0$, assuming \hat{w}_n is ergodic. For the model (1), w_n is rarely uncorrelated, meaning this test is not valid. For the model (5), this whiteness test is only valid when \hat{w}_n consistently estimates w_n . If we only consistently estimate the polynomials $A(z)$ and $B(z)$ needed for OAR, we cannot use this method either.

Jervis in [36] did present a quantitative method for assessing OAR methods. This method, however, required the subject to perform periodic vertical or horizontal eye movements, which produced rectangular waveforms in the EOG. The autocorrelation function of periodic rectangular waves is a periodic sequence of triangular waves. Measuring the residual of this triangular waveform in the autocorrelation function of the corrected EEG provides a performance index of the OAR procedure. This quantitative measure only works for data acquisition systems that do not significantly distort the rectangular waves produced by vertical eye movements, and assumes that natural eye movements produce rectangular waves. Unfortunately, we could not use this measure in this paper because the anti-aliasing filters of our data acquisition system did appreciably distort these rectangular waveforms.

In this paper, we use a combination of visual inspection and minimization of the cumulative error (31) to evaluate OAR performance.

3. OAR Examples. In this section we present examples to compare the AR-MAX and linear regression OAR subtraction methods. Our first example analyzes a 20 second segment of data that begins with 5 voluntary blinks followed by a series of horizontal and vertical eye movements. We fit a (10,10,10) model to the data using 24 channels of measured EEG as y_n and 4 channels of measured EOG as u_n . We did not use all 7 channels shown in Figure 2(b) as EOG inputs, but instead used only the lower four as EOG and the upper 3 as EEG. We did so for three reasons. First, channels FP1, FPz, and FP2 had more ocular artifact than the other EEG channels, and therefore provide a challenging “test-bed” for comparing performance of the removal algorithms and allow reliance on “gold-standard” visual inspection of results to judge performance. Second, Jervis *et al.* in [5] did not have good results using these lower 4 electrodes in the regression OAR subtraction method, thereby allowing the work herein to demonstrate an improvement upon their results. Finally, Jervis *et al.* in [18]

recommends using the fewest number of channels located above the eyes to reduce the correlation between EEG and EOG, in an attempt to ensure that assumption **A3** is satisfied.

Figure 3 shows the 20 seconds of measured EOG and EEG data referenced to the mastoids. The four EOG channels appear first followed by the 24 EEG channels. Figure 4 shows the estimated background EEG using the ELS estimator with $\lambda = 1 = a_n \forall n$. There was a significant decrease in the magnitudes of AR and MA coefficient estimates beyond first order, making it appear that a (1,1,1) model is sufficient for OAR. We will see later, however, that when we perform OAR on a single EEG channel, a (1,1,1) model leads to poor performance. We also observed that the matrices for the autoregressive and moving average terms are approximately diagonal, indicating that each channel of measured EEG's AR and MA terms are primarily a function of previous values of the same channel and not those from other channels. This observation suggests that each EEG channel should be used separately to perform OAR, thereby significantly reducing the required computations and variance of the estimator.

We next applied the order estimation procedure described in Section 1.5 using $\lambda = 1 = a_n \forall n$ on a 100 second section of data. Figure 5 shows the *CIC* values, cumulative errors (the first term in (30)), and order costs (the second term in (30)) for each order triplet $(k, k, k), k \in \{0, 1, \dots, 11\}$. The algorithm chose a maximum model order of 3, and then proceeded to determine individually the orders of the autoregressive, exogenous, and moving average terms, arriving at a selected order of (3,3,2). In general, as the model order increased, the cumulative error decreased, with the largest percentage decrease occurring when the autoregressive order increased from zero to one. Other than this sharp decrease, the order estimation procedure did not produce clear results. Beyond the addition of an autoregressive term, the choice of (3,3,2) seems somewhat arbitrary because the *CIC* curves for moving average and exogenous orders did not appear to have definitive minimums.

As previously mentioned, another way to empirically determine model order is to over-parametrize the model and choose the highest model order that produces non-zero coefficient estimates. We systematically tested ARMAX models of increasing order on a 300 second segment of data containing significant eye movements, both voluntary and involuntary. Figures 6-7 show the last 100 seconds of the four EOG channels (LOC, LIO, RIO, ROC), the raw ocular artifact corrupted EEG recorded on the FP1 channel and the results of ocular artifact removal for various models/orders. Figure 6(b) illustrates the improvement obtained when using a forgetting factor to weight the data and thereby track potentially time-varying coefficients with ELS as compared to the unweighted results shown in Figure 6(a). A forgetting factor with a 2 second half-life was chosen, accounting for the timescale of the artifacts. Figures 7(a),(b),(c) show that failure to incorporate an autoregressive term in the

model results in poor performance. Without an autoregressive term, the OAR performance does not improve but, in fact, even degrades as the orders of exogenous and/or moving average terms are increased. These data suggest that increasing the order of exogenous or moving average terms by themselves does not improve performance (see Figures 7(a),(b)) but increasing these orders in combination with that of the autoregressive term (as seen in Figures 6(b) and 7(c)) does result in performance improvements up until order four or five. Increases of order beyond these levels can actually degrade performance.

4. Conclusions. In this paper, we expanded the zero order exogenous regression model (1) traditionally used for OAR into a general autoregressive (AR), moving average (MA), and exogenous (X) model (5). Using the extended least squares algorithm with a two second half-life exponential weighting of past data, removed some ocular artifact that the regression model could not. On the other hand, the regression model better removed other types of artifacts such as blinks. In general, we observed that some models removed certain types of artifact quite well, while at the same time they removed other types poorly.

Furthermore, low order ARMAX models did not remove artifact as effectively as the regression model, possibly because a bias developed due to the undermodeling of the background EEG. This bias would cause the improper subtraction of the artifact and even might introduce artifact not present in the original EEG. This bias was not present in the regression model method, because undermodeling the background EEG in this model results in a suboptimal, yet unbiased, estimator.

We also examined an ARMAX order estimation algorithm based on the CIC information criterion. This algorithm produced third order models, yet we saw that these third order models did not necessarily remove ocular artifact as well as the regression model or fourth/fifth order ARMAX models. We also observed that increasing the model order did not necessarily produce better ocular artifact removal.

In addition, we saw that the introduction of a forgetting factor in the recursive least squares estimator for all model orders removed ocular artifact better than assuming a time-invariant model. This forgetting factor was essential for good performance of the ARMAX model OAR.

In summary, high order ARMAX models can sometimes remove ocular artifact that regression models cannot; however, they do so at the expense of increased algorithm complexity, and the risk of introducing new ocular artifact if they do not properly model the underlying background EEG. Regression models, while being suboptimal, are simpler to implement and do not incur this undesirable risk. The regression models can also remove some types of artifact better than the ARMAX models.

Acknowledgements. The authors would like to thank Jeff Ingram for his help collecting the data for this paper and many fascinating and challenging projects during these past two years. We also want to thank Tyrone Duncan, and Bill Paschke for their valuable feedback.

REFERENCES

- [1] S. A. HILLYARD, *Methodological issues in CNV research*, in: Bioelectric recording techniques part B. electroencephalography and human brain potentials, M. M. Patterson R. F. Thompson, Ed., pp. 259–280. Academic Press, 1974.
- [2] L. A. GEDDES AND L.E. BAKER, *Principles of Applied Biomedical Instrumentation*, Wiley & Sons, New York, 1975.
- [3] R. VERLEGER, T. GASSER, AND J. MOCKS, *Correction of EOG artifacts in event-related potentials of the EEG: aspects of reliability and validity*, *Psychophysiology*, 19(1982), pp. 472–480.
- [4] V. ROWLAND, *Cortical steady potential (direct current potential) in reinforcement and learning*, in: Progress in physiological psychology, J. M. Sprague E. Stellar, Ed., Vol. 2, pp. 1–77. Academic Press, New York, 1968.
- [5] B. W. JERVIS, E. C. IFEACHOR, AND E. M. ALLEN, *The removal of ocular artefacts from electroencephalogram: A review*, *Medical & Biological Engineering & Computing*, 26(1988), pp. 2–12.
- [6] P. K. SADASIVAN AND D. NARAYANA DUTT, *ANC schemes for the enhancement of EEG signals in the presence of EOG artifacts*, *Computers and Biomedical Research*, 29(1996), pp. 27–40, Article No. 0003.
- [7] E. C. IFEACHOR AND B. W. JERVIS, *Digital Signal Processing: A Practical Approach*, Addison-Wesley Publishers Ltd., 1993.
- [8] G. GRATTON AND M. G. H. COLES, *Generalization and evaluation of eye-movement correction procedures*, in: Correcting ocular artifacts in the EEG: a comparison of several methods, C. H. Brunia *et al.*, Ed., Vol. 3, pp. 14–16. *Journal of Psychophysiology*, 1989.
- [9] T. ELBERT, W. LUTZENBERGER, AND N. BIRBAUMER, *Removal of ocular artifacts from the EEG - a biophysical approach*, *Electroencephalography and Clinical Neurophysiology*, 60(1985), pp. 455–463.
- [10] W. LUTZENBERGER AND T. ELBERT, *An attempt of EOG correction*, in: Correcting ocular artifacts in the EEG: a comparison of several methods, C. H. Brunia *et al.*, Ed., Vol. 3, pp. 20–21. *Journal of Psychophysiology*, 1989.
- [11] M. M. C. VAN DEN BERG-LENSSEN AND C. H. M. BRUNIA, *Correction of ocular artifacts in EEGs using an autoregressive model to describe the EEG*, in: Correcting ocular artifacts in the EEG: a comparison of several methods, C. H. M. Brunia *et al.*, Ed., Vol. 3, pp. 26–29. *Journal of Psychophysiology*, 1989.
- [12] G. VAN DRIEL, J. C. WOESTENBURG, AND A. W. VAN BLOKLAND-VOGELESANG, *Frequency domain methods: A solution for the problems of EOG-EEG contaminations in ERPs*, in: Correcting ocular artifacts in the EEG: a comparison of several methods, C. H. Brunia *et al.*, Ed., Vol. 3, pp. 29–34. *Journal of Psychophysiology*, 1989.
- [13] P. R. KUMAR AND P. VARAIYA, *Stochastic Systems: Estimation, Identification, and Adaptive Control*, Prentice-Hall, Inc., 1986.
- [14] M. H. A. DAVIS AND R. B. VINTER, *Stochastic Modelling and Control*, Chapman and Hall, 1985.
- [15] E. C. IFEACHOR, *Investigation of Ocular Artefacts in the Human EEG and Their Removal by a Microprocessor-based Instrument*, Ph.D. thesis, Plymouth Polytechnic, Dec. 1984.

- [16] F. H. LOPES DA SILVA, *Analysis of EEG non-stationarities*, Contemporary Clinical Neurophysiology, pp. 163–179, 1978, EEG Supplement No. 34.
- [17] H. J. F. DE WIT, *Testing of stationarity of background EEG and a single trial analysis using orthogonal polynomials*, M. S. thesis, Eindhoven University of Technology, 1985.
- [18] B. W. JERVIS, E. C. IFEACHOR, AND M. COELHO, *The removal of ocular artifacts from the EEG*, in: Correcting ocular artifacts in the EEG: a comparison of several methods, C.H. Brunia *et al.*, Ed., Vol. 3, pp. 16–20. Journal of Psychophysiology, 1989.
- [19] P. D. TUAN, *A comment from the viewpoint of time series analysis*, in: Correcting ocular artifacts in the EEG: a comparison of several methods, Cornelis H. Brunia *et al.*, Ed., Vol. 3, pp. 46–48. Journal of Psychophysiology, 1989.
- [20] N. R. DRAPER AND H. SMITH, *Applied Regression Analysis*, Wiley & Sons, New York, 1966.
- [21] G. GRATTON, M. G. H. COLES, AND E. DONCHIN, *A new method for off-line removal of ocular artifacts*, Electroencephalographic & Clinical Neurophysiology, 55(1983), pp. 468–484.
- [22] P. MANDL, T. E. DUNCAN, AND B. PASIK-DUNCAN, *On the consistency of a least squares identification procedure*, Kybernetika, 24:5(1988), pp. 340–346.
- [23] V. SOLO AND X. KONG, *Adaptive Signal Processing Algorithms: Stability and Performance*, Prentice Hall, New Jersey, 1995.
- [24] B. BERCU, *Weighted estimation and tracking for ARMAX models*, SIAM Journal on Control and Optimization, 33:1(1995), pp. 89–106.
- [25] B. WIDROW, J. MCCOOL, AND M. BALL, *The complex LMS algorithm*, Proceedings of the IEEE, pp. 719–720, April 1975.
- [26] H. CHEN AND L. GUO, *Identification and Stochastic Adaptive Control*, Birkhauser, 1991.
- [27] P. EYKHOFF, *System Identification*, Wiley & Sons, 1974.
- [28] S. M. HAAS, *Decision weighted adaptive algorithms with applications to wireless channel estimation*, M. S. thesis, University of Kansas, April 1999.
- [29] H. CHEN AND L. GUO, *Convergence rate of least-squares identification and adaptive control for stochastic systems*, International Journal on Control, 44:5(1986), pp. 1459–1476.
- [30] L. GUO, *Self-convergence of weighted least-squares with applications to stochastic adaptive control*, IEEE Transactions on Automatic Control, 41:1(1996), pp. 79–89.
- [31] H. CHEN, *Recursive Estimation and Control for Stochastic Systems*, Wiley & Sons, 1985.
- [32] E. J. HANNAN AND J. RISSANEN, *Recursive estimation of mixed autoregressive-moving average order*, Biometrika, 69:1(1982), pp. 81–94.
- [33] D. Q. MAYNE AND F. FIROOZAN, *Linear identification of ARMA processes*, Automatica, 18:4(1982), pp. 461–466.
- [34] E. J. HANNAN AND L. KAVALIERIS, *Linear estimation of ARMA processes*, Automatica, 19:4(1983), pp. 447–448.
- [35] B. W. JERVIS, M. J. NICHOLS, E. M. ALLEN, N. R. HUDSON, AND T. E. JOHNSON, *The quantitative assessment of electroencephalograms corrected for eye movement artefact*, in: 1st European Signal Processing Conference, Lausanne, Switzerland, September 1980.
- [36] B. W. JERVIS, M. J. NICHOLS, E. M. ALLEN, N. R. HUDSON, AND T. E. JOHNSON, *The assessment of two methods for removing eye movement artefact from the EEG*, Electroencephalographic & Clinical Neurophysiology, 61(1985), pp. 444–452.

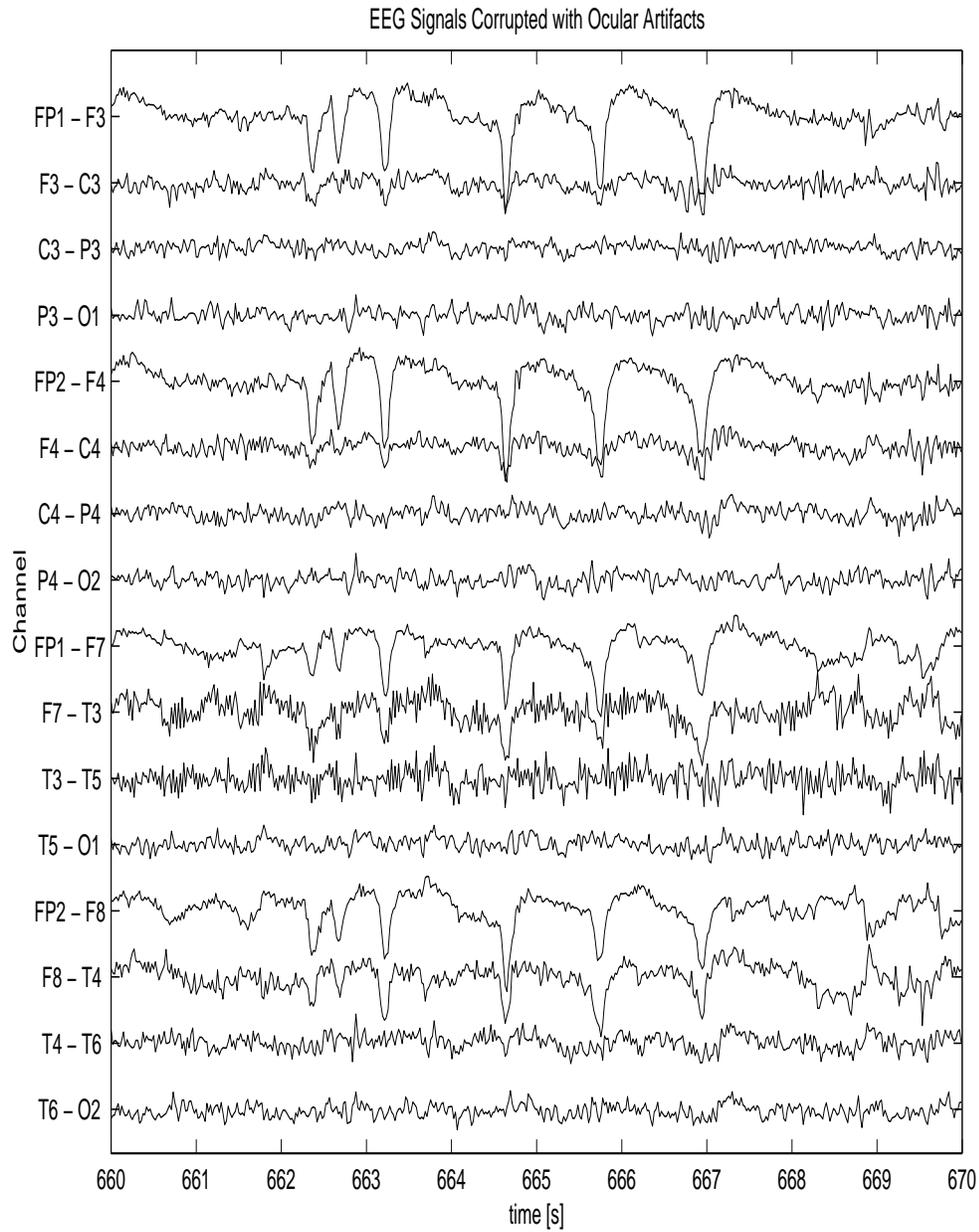
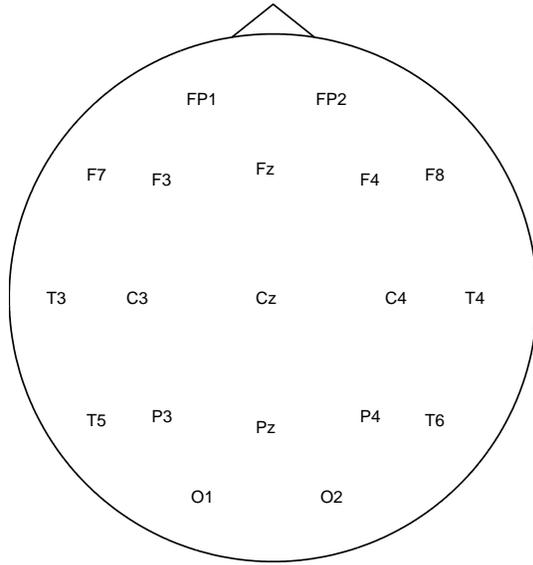


FIG. 1. A segment of EEG signals corrupted with ocular artifact

(a)



(b)

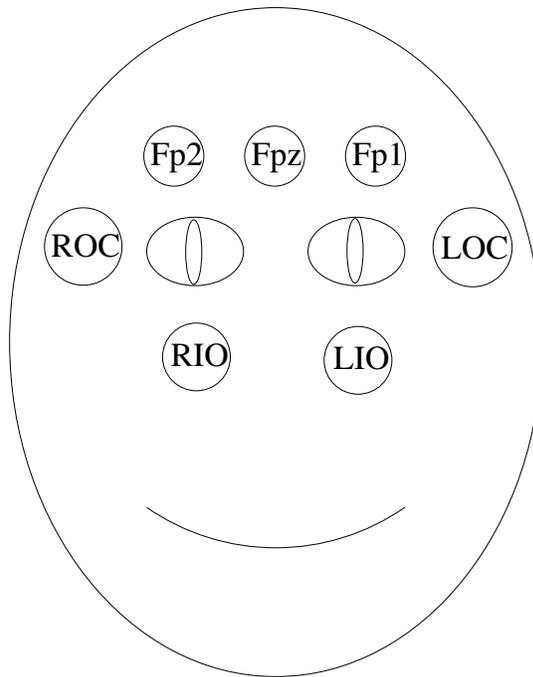


FIG. 2. Location of standard electrode placement for (a) scalp recorded electroencephalogram (EEG) (top view) and (b) electrooculogram (EOG)

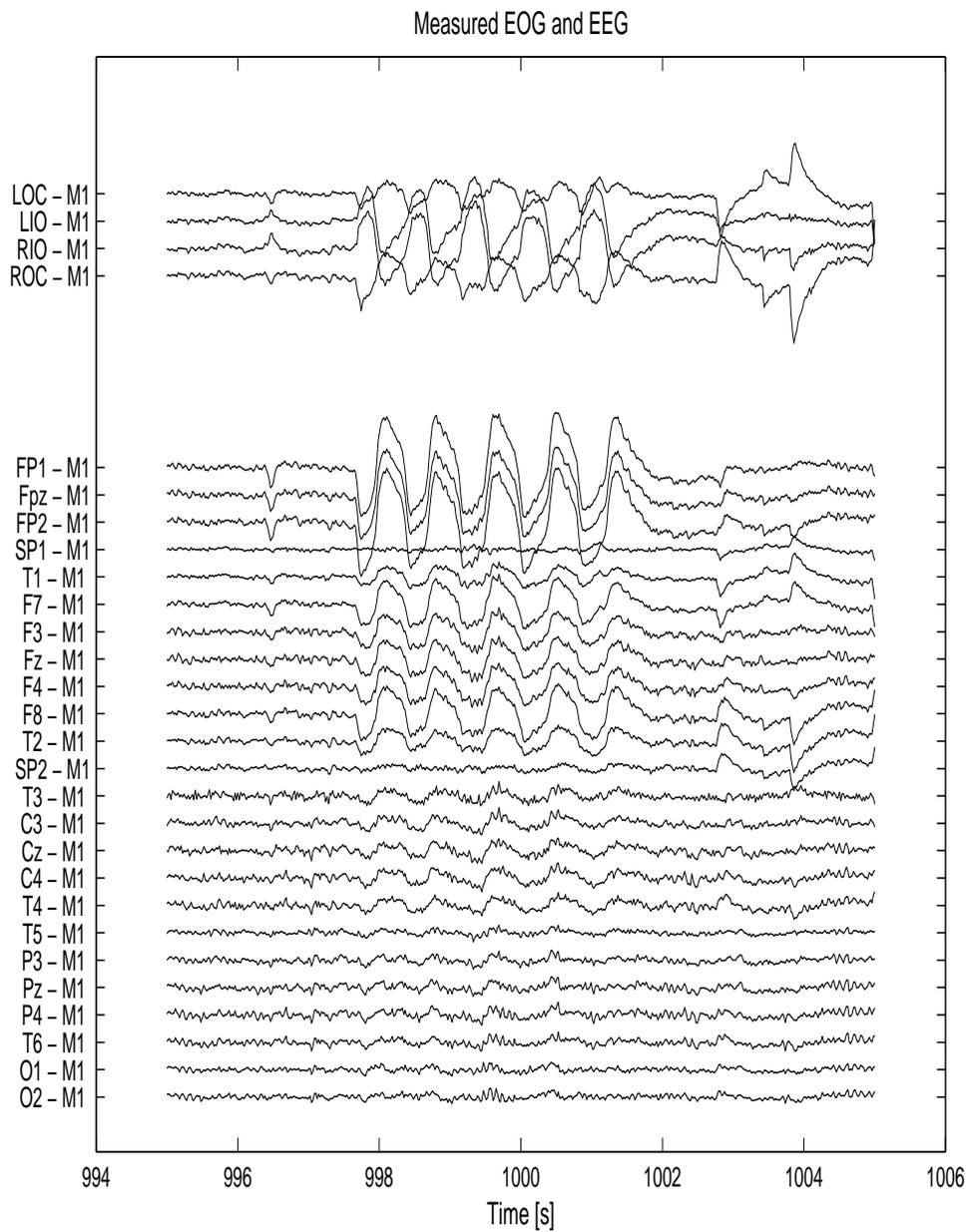


FIG. 3. The twenty second segment of data used to train a $(10,10,10)$ ARMAX model for ocular artifact removal

Measured EOG and Estimate of Background EEG Using Final Parameters (ARMAX Order: [10,10,10])

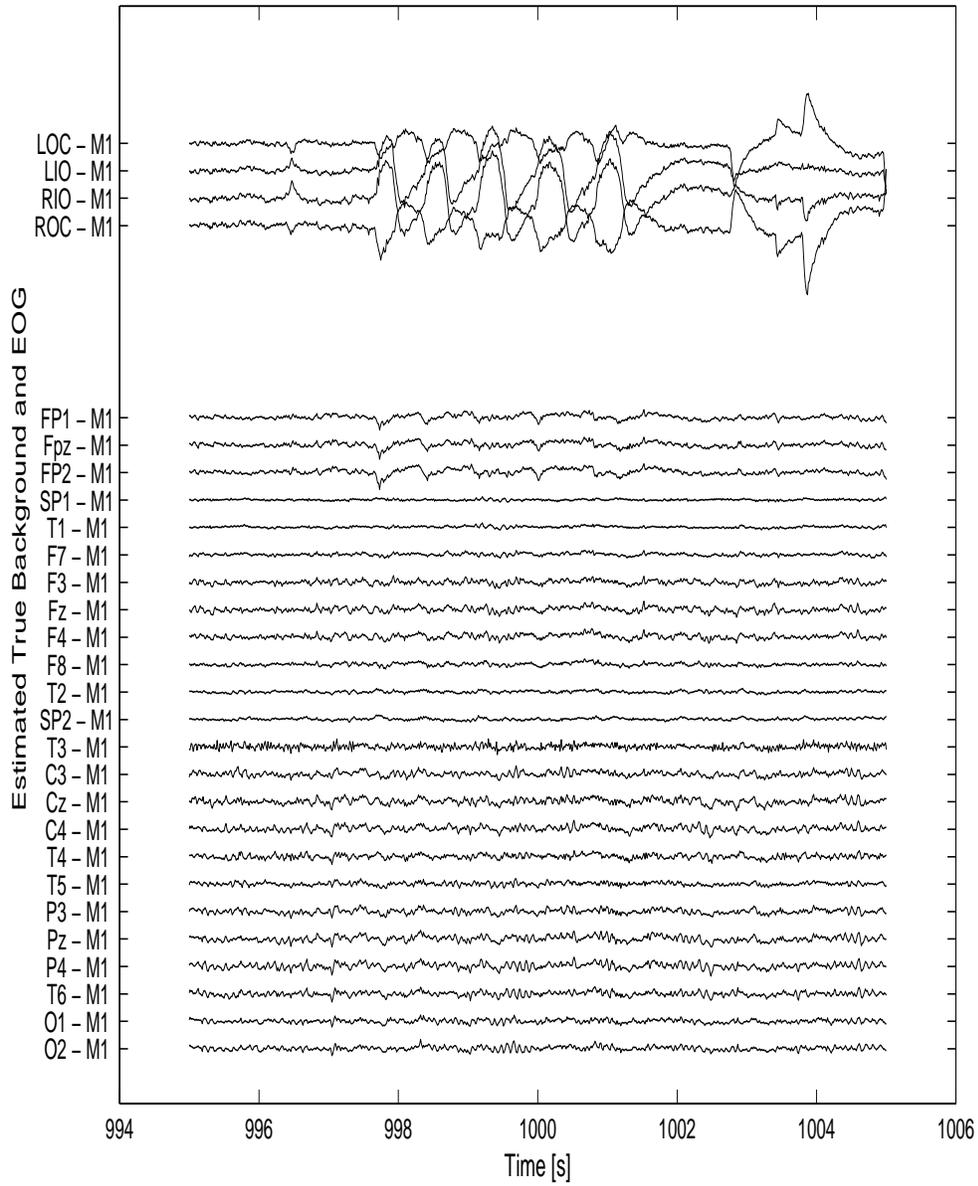


FIG. 4. Estimated background EEG using (10,10,10) ARMAX OAR

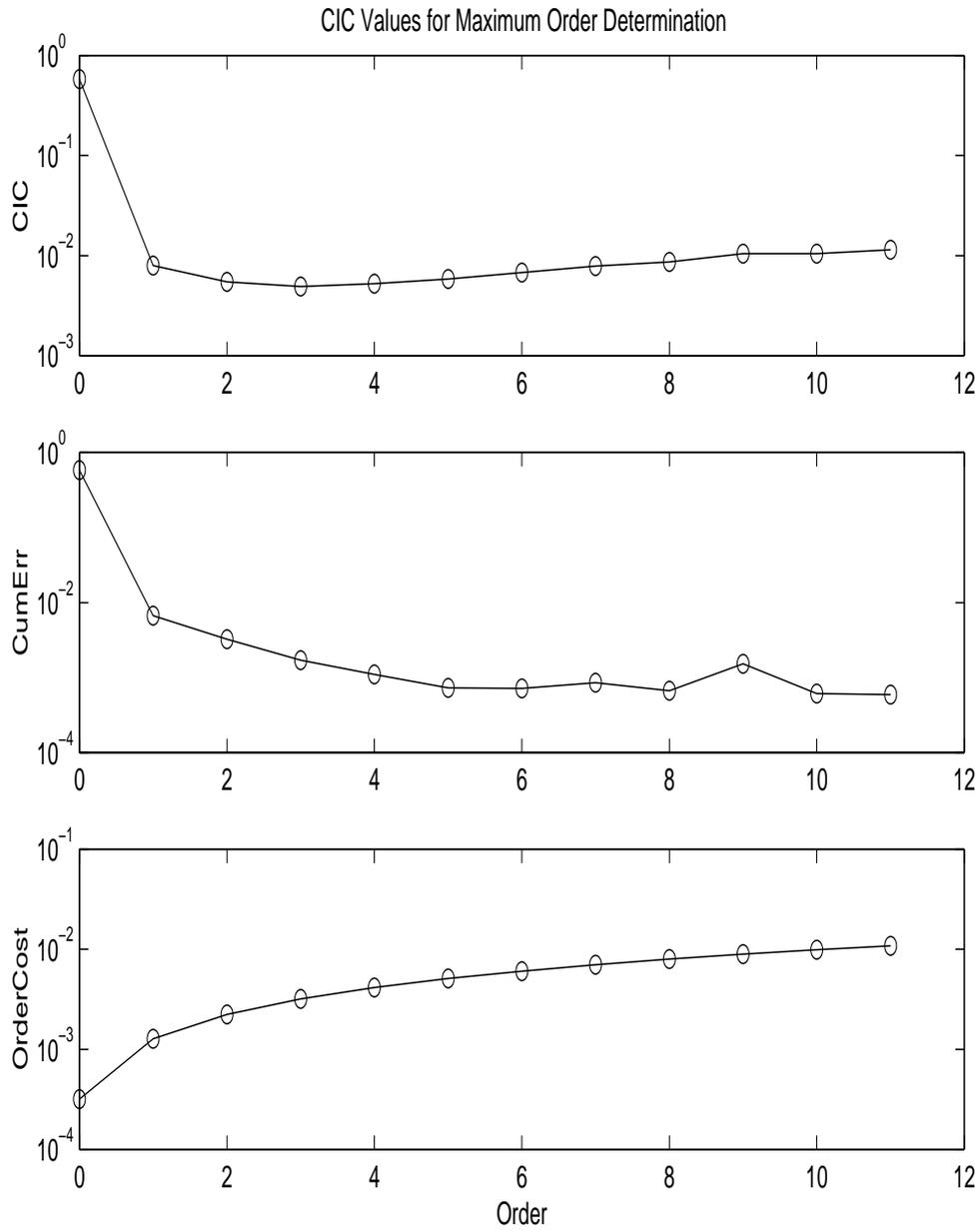
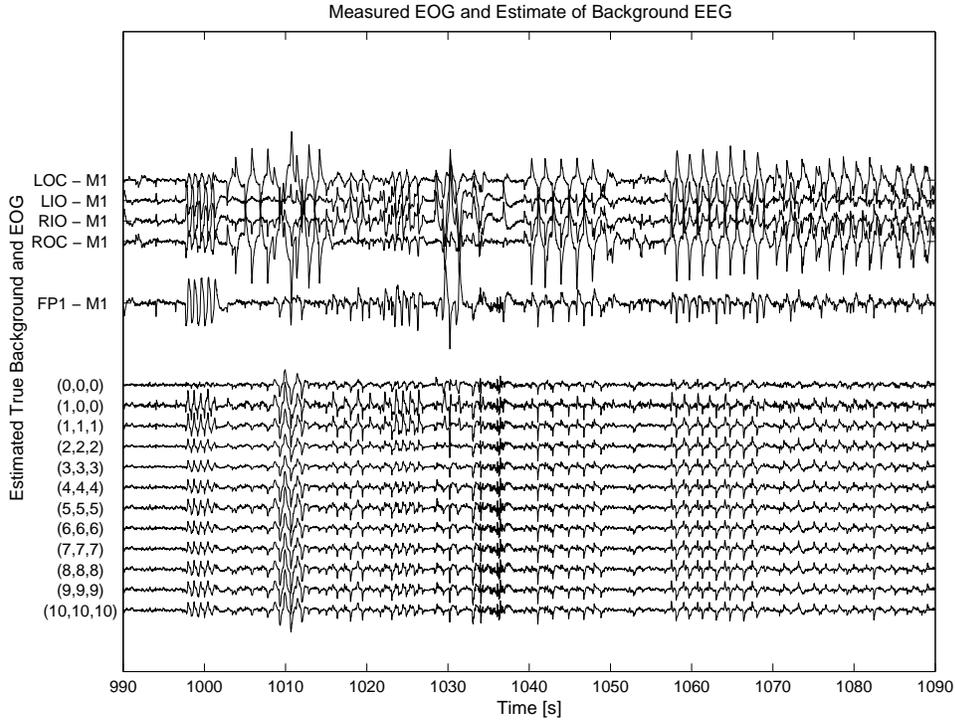


FIG. 5. Estimating the maximum order of an ARMAX model fit to the EEG channel FP1

(a)



(b)

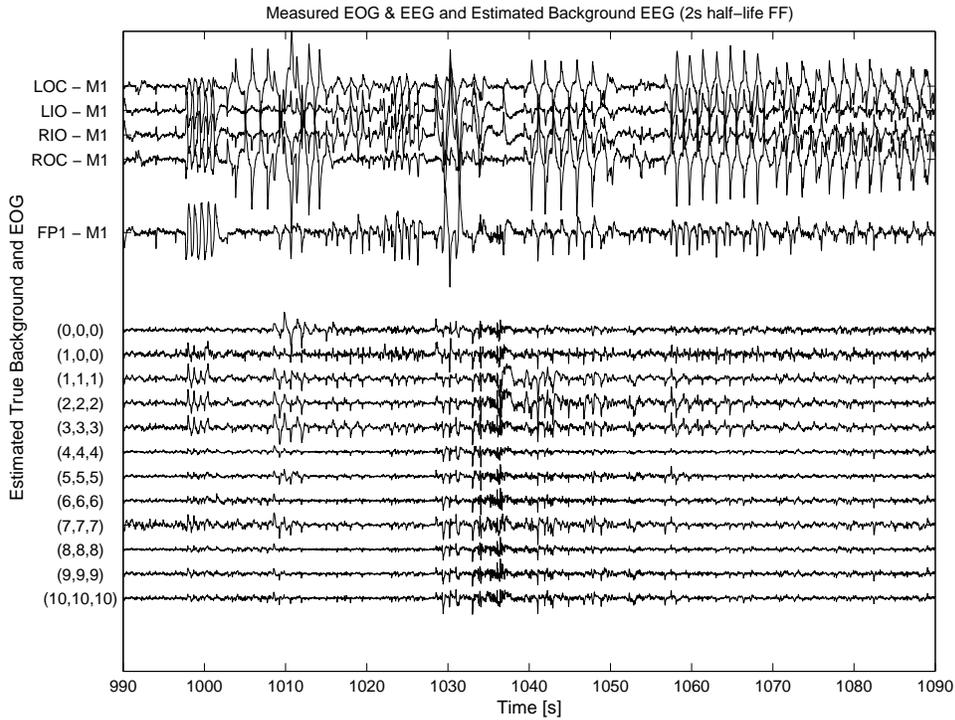
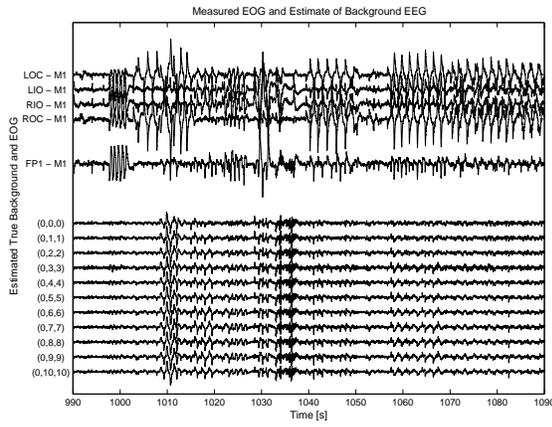
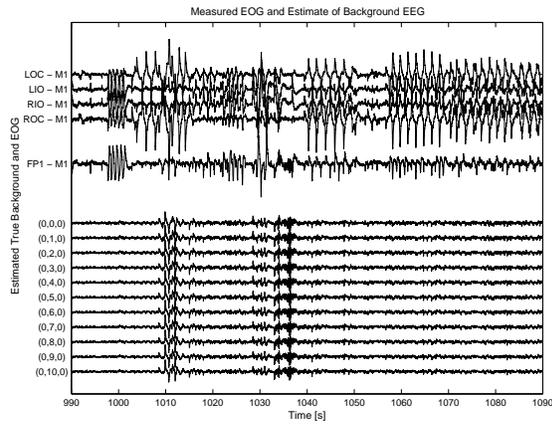


FIG. 6. ARMAX OAR examples with (a) $\lambda = 1 = a_n \forall n$, (b) $\lambda = 0.9986$ (2 second half-life @240 Hz) and $a_n = 1 \forall n$.

(a)



(b)



(c)

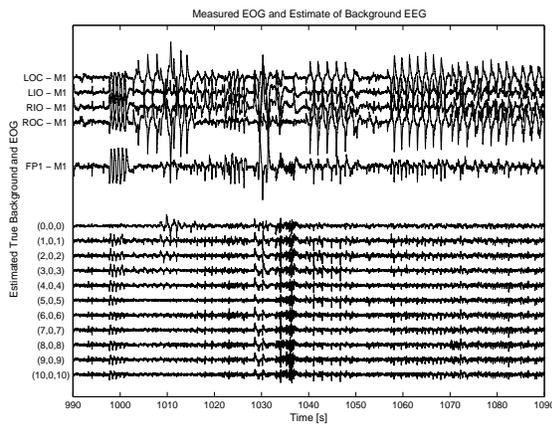


FIG. 7. Examples illustrating (a) XMA models, (b) Exogenous models, and (c) ARMAX models with zero order exogenous terms. All models have $\lambda = 1 = a_n \forall n$, (b) $\lambda = 0.9986$ (2 second half-life @240 Hz) and $a_n = 1 \forall n$.