

Eye-Gaze and Arm-Movement Interaction

Ritaja Sur¹, Lior Noy², Tamar Flash², Benjamin Kedem¹

¹University of Maryland, College Park, MD 20742, USA
ritaja@math.umd.edu, bnk@math.umd.edu

²Weizmann Institute of Science, Rehovot 76100, Israel
lior.noy@weizmann.ac.il, tamar.flash@weizmann.ac.il

September 2010

Abstract

Eye-gaze in response to arm movement stimuli is modeled. It is demonstrated that the eye-gaze response depends to a large degree on both instantaneous and past arm components movement as well as on past “long memory” of eye-gaze on a scale of about 0.1 second. The suggested class of models provide close fits. A long memory property in eye-gaze is validated from a certain spectral condition.

1 Introduction

To reorient eye-gaze, human scene perception undergoes rapid eye movements, called saccades, occurring at the rate of about three times per second. Given dynamic stimuli, such as arm movement, this translates into one-dimensional x- and y-coordinate eye-gaze time series characterized by rapid oscillation fraught with sudden discontinuities, and distinguished by a spectrum heavily supported at frequency zero. The later property induces “long memory” in time, the central issue of this paper. Figure 1 depicts a typical x-coordinate eye-gaze time series response and corresponding input from arm components.

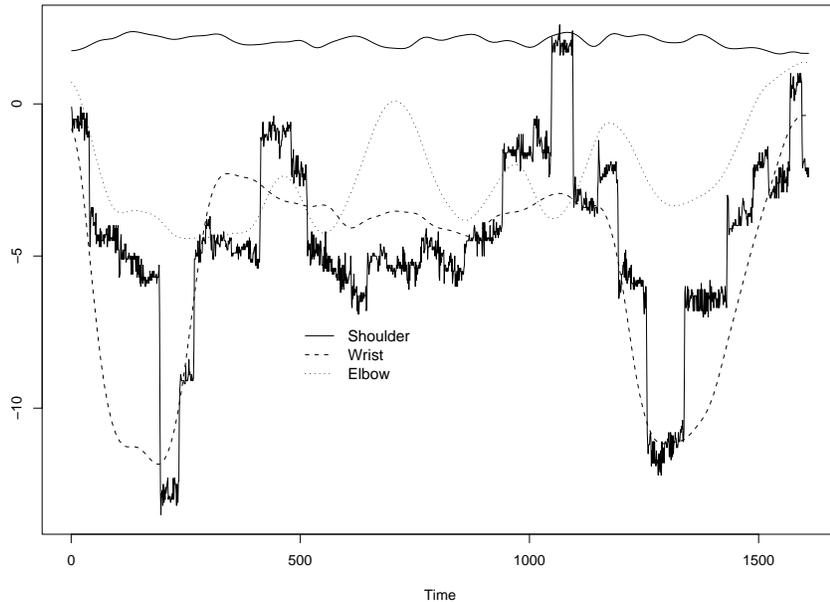


Figure 1: A time series from subject 5, movement 9, x-coordinate, under “Watch”, and arm inputs sampled at 120Hz.

It is intuitively clear that in order to predict the future it is necessary incorporate past information. The question is then how far into the past do we to go to obtain meaningful prediction. This question is considered here in terms of eye-gaze and arm movement relationship. More precisely, the goal of this paper is to quantify the memory induced by eye-gaze stimulated by arm movement, and to describe the influence of certain past and present covariates on eye-gaze as a function of time. This is done by fitting to the x-, y-, and $\sqrt{x^2 + y^2}$ -coordinates time series linear prediction models containing past eye-gaze as well as past and instantaneous inputs from arm components movement. The statistical analysis, illustrated by a few typical cases, reveals a remarkable regularity and consistency of eye-gaze/arm-movement relationship across seven subjects and ten arm-movements under two conditions. The analysis points to: a. Remarkably close fits with well behaved residuals of models where the eye-gaze response depends on arm inputs as well as on past response. b. The fact that past memory on the order of about 0.1 second is important for prediction,

but that the immediate past on the order of 16-20 milliseconds (roughly) is crucial. c. The existence of long memory traits as defined by spectral considerations.

1.1 Some Preliminaries

A *time series* or *process* is a collection of data points $y_t, t = 1, 2, \dots, n$ measured at successive time points. Due to natural ordering in time, adjacent data points in a time series tend to be correlated. Hence, the notion of autoregression where the present value at time y_t depends on past values at previous time points y_{t-1}, y_{t-2}, \dots is natural. In many cases, in addition to past values, a time series may depend on other series as well. By a *stationary* time series or process we shall mean a series for which the correlation between two values y_{t_1} and y_{t_2} depends on the distance between them $|t_2 - t_1|$. A well known model useful for short memory stationary processes is the autoregressive moving average model (ARMA). The so called autoregressive fractionally integrated moving average model (ARFIMA) is useful for a certain class of long memory processes [1], [15]. A precise definition of “long memory” is given below.

In general, the dependence structure of a stationary time series is described by the *autocovariance* or by the *autocorrelation* function which measure the covariance and correlation between pairs of observations, respectively. The autocovariance function γ is a function of the time lag h ,

$$\gamma(h) = Cov(y_t, y_{t+h})$$

and the autocorrelation is given by the normalized quantity $\gamma(h)/\gamma(0)$. When it exists, the Fourier transform of the sequence $\gamma(h)$ is a function of frequency referred to as the *spectral density* and denoted by $f(\lambda)$. The spectral density describes the power or variance in a signal as a function of frequency, and a process for which the power is distributed evenly across all frequencies (flat spectrum) is referred to as *white noise*. An important example of white noise is a sequence of uncorrelated observations with mean zero and fixed variance. The quality of a regression model is judged by the closeness of its residuals to white noise.

The sample analog of the autocovariance is the *sample autocovariance*,

$$\hat{\gamma}(h) = \frac{1}{n} \sum_t (y_t - \bar{y})(y_{t-h} - \bar{y}).$$

To test for whiteness, one checks whether the *sample autocorrelation* $\hat{\gamma}(h)/\hat{\gamma}(0)$, $h = 1, 2, \dots$ is sufficiently small. As an example, the sample autocorrelation shown at the top of Figure 6 points to a series which is far from white noise, whereas the bottom part shows a sample autocorrelation typical of white noise. Another useful sample quantity is the *periodogram* defined as

$$I(\lambda) = \frac{1}{n} \left| \sum_t e^{-i\lambda t} y_t \right|^2.$$

Summing the periodogram over frequency gives the *cumulative periodogram* (cpgram), an important diagnostic tool. A typical cumulative periodogram from white noise is close to a 45-degree line, an example of which is shown in Figure 7 on the right. Marked deviations

from the 45-degree line, as in Figure 7 on the left, point to series far removed from white noise.

Long memory time series refer to processes where observations far away from each other are strongly correlated. In the case of ARMA processes, the covariance between Y_t and Y_{t+k} decays exponentially fast as $k \rightarrow \infty$, and is absolutely summable under general conditions. For this reason ARMA processes are referred to as short memory processes. A process for which

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty$$

is said to be a long memory process [15]. There are other definitions of long memory processes of which we shall adopt a somewhat narrower definition in the spectral domain. Accordingly, a long memory process is a stationary process for which the spectral density function $f(\lambda)$ satisfies the condition

$$\lim_{\lambda \rightarrow 0} \frac{f(\lambda)}{c |\lambda|^{-2d}} = 1 \quad (1)$$

for *memory parameter* d such that $d \in (0, .5)$ and a positive constant c . Condition (1) implies that the spectral density has a pole at frequency zero.

Of the various methods available for the estimation of the memory parameter d we shall adopt Whittle’s estimation method [18]. Details of this estimation procedure can be found in [7], [5], [8].

The study of long memory characteristics in time series has long been under investigation in various fields including hydrology, economics and biology. Many examples of real-life time series which possess this long memory behavior can be found in [2], [3], [17]. In relation to the present work, experiments show that the human sensorimotor coordination system falls into the framework of long memory processes [4].

1.2 Data Description

Employing an eye-gaze tracker, eye-gaze data were recorded for 7 individuals, referred to as subjects, in response to 10 distinct simple one-arm movements of a demonstrator [14]. The eye-gaze was recorded under two conditions. Under the first condition the subjects were instructed to watch the movements and answer a few questions regarding the movements. Under the second condition the subjects were asked to watch and then imitate the arm movements as closely as possible. The first condition is referred to as “Watch” and the second as “Imitate”. The x-y coordinates of the wrist, elbow, and shoulder of the demonstrator, and those of the eye-gaze of the subjects were recorded giving $7 \times 10 \times 2 = 140$ data sets. The eye-gaze data analyzed here consist of x-coordinate time series sampled at 120Hz and ranging in length from 801 (6.68s) to 1609 (13.41s) data points. The arm data were taken from a video recorded at 15Hz. More details of the experimental apparatus, procedure and data analysis used to derive the eye-gaze coordinates can be obtained from the work of Noy [14].

2 Data Analysis

Since the arm movement serves as an input to eye-gaze output, it is clear that eye-gaze and arm movement must be related. To explore this relationship, we start with a simple model.

We model the x-coordinate of eye-gaze time series denoted, by y_t , where the first 50 time points were truncated to eliminate transient effects at the beginning of each signal. For this purpose, consider first a simple model referred to as model M0, depending only on the x-coordinates of wrist (w_t), elbow (e_t), and shoulder (sh_t) as regressors or covariates,

$$(M0) \quad y_t = \beta_0 + \beta_1 w_t + \beta_2 e_t + \beta_3 sh_t + \epsilon_t. \quad (2)$$

where ϵ_t is noise or an error term. Following [10], the coefficient were estimated using least squares fitting. Across the two conditions, 10 movements, and 7 subjects, it was observed that the wrist component coefficient β_1 was significant at 5% level for almost all cases (with exception of three cases). For coefficients β_2 and β_3 , there were more than three cases where the coefficients were insignificant at 5% level. Let us consider a representative case, say Subject 4 movement 1 under the “Watch” condition. We performed the least squares fit and obtained goodness-of-fit statistics and residuals. Figure 2 shows the least squares fit of the model. The sample autocorrelation function of the residuals resulting from the linear regression fit is shown in Figure 3. The autocorrelation plot clearly shows a slow decay, pointing to significant values even at lag greater than 30. In addition, and this is at the core of the present work, the spectral density computed for the residuals shows a markedly large peak around zero frequency in Figure 4, indicating the possibility of long memory behavior in this time series. This behavior was observed in all cases. To validate the presence of long memory in the sense of (1) in the residual series, the fractional parameter d was estimated using Whittle’s method [8]. The estimated d averaged for each subject across 10 movements for both conditions are reported in Table 1. All but one average d value was found to be in the interval $(0, 0.48)$, pointing to substantial long memory. These results imply that a significant long memory structure has not been taken into account in model M0 in (2). This suggests adding regressors pertaining to past response to improve the fit. As we shall see next, augmenting model M0 with past eye-gaze response improves the fit greatly.

To account for the long memory residuals in model M0 in (2), we appeal to the representation of a long memory process by an infinite autoregressive (AR) expansion [15], [9]. Since the length of each eye-gaze time series is finite, we fitted an $AR(\tilde{p})$ with a finite order \tilde{p} to the estimated regression residual process $\hat{\epsilon}_t$. The new residuals obtained by fitting a long AR (\tilde{p}) did not show any long range dependence. That is, the autocorrelation plot, the cumulative periodogram plot along with Portmanteau tests confirmed the whiteness of these residuals for all cases.

This analysis suggests models where the instantaneous eye-gaze depends on past eye-gaze as well as on present and past movement of arm components. In the next section we provide several examples of such models and their quality of fit.

3 Prediction Models for Eye-gaze

In light of the previous discussion, we explore eye-gaze/movement relationship by including appropriate past and present covariates, to obtain fine-tuned models which trace the gaze

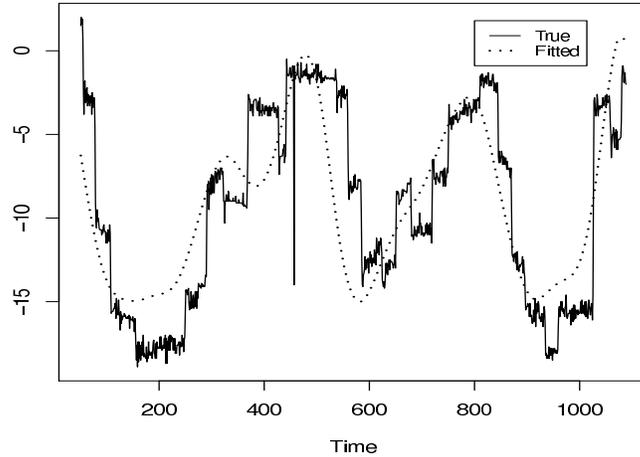


Figure 2: True and fitted values for Subject 4, Movement 1, “Watch” condition.

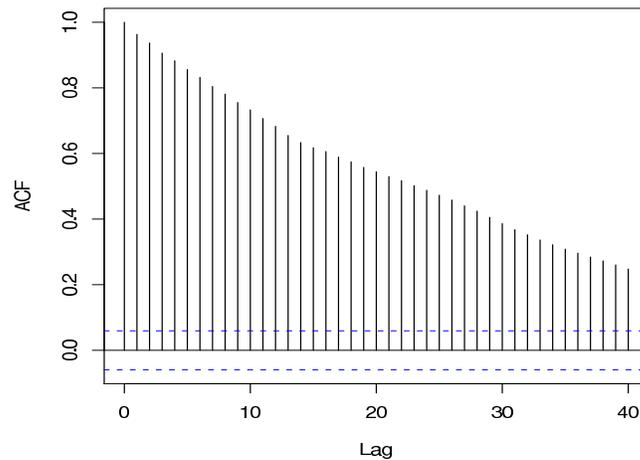


Figure 3: Autocorrelation from M0 residuals. Subject 4, Movement 1, “Watch” condition. The dashed lines show the 95% confidence limits based on white noise.

data in detail. We have examined numerous cases all of which behave quite similarly, across subjects, movements, and conditions, and thus it suffices to illustrate our findings by a few typical cases of x-coordinate, y-coordinate, and and r-coordinate ($r = \sqrt{x^2 + y^2}$) series. These cases are:

- Subject 5, Watch, Movement 9, x-coordinate (S5W9x),
- Subject 3, Watch, Movement 8, x-coordinate (S3W8x),
- Subject 3, Watch, Movement 6, y-coordinate (S3W6y),
- Subject 2, Watch, Movement 4, r-coordinate (S2W4r),
- Subject 4, Imitate, Movement 7, r-coordinate (S4I7x).

Regardless of coordinate, the gaze, wrist, elbow, and shoulder time series are denoted by y_t ,

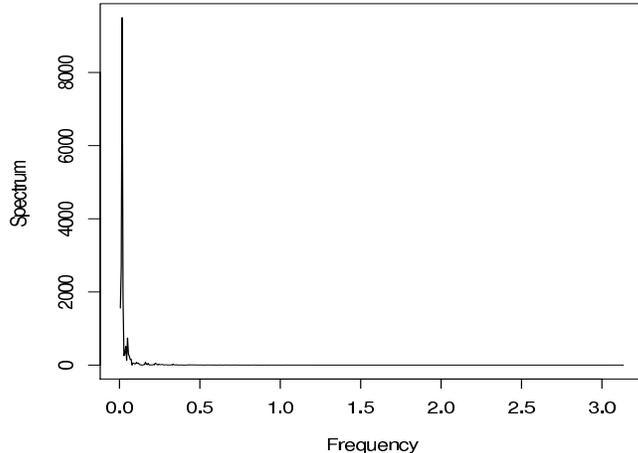


Figure 4: Estimated spectral density from the M0 residuals. Subject 4, Movement 1, “Watch” condition.

Table 1: Average d estimates from residuals of M0 using Whittle’s method with standard deviations shown in parenthesis.

Subject	Watch	Imitate
1	0.29 (0.05)	0.28 (0.09)
2	0.34 (0.16)	0.36 (0.08)
3	0.32 (0.11)	0.35 (0.10)
4	0.26 (0.10)	0.31 (0.07)
5	0.47 (0.08)	0.46 (0.09)
6	0.17 (0.09)	0.29 (0.10)
7	0.44 (0.09)	0.56 (0.13)

w_t , e_t , and s_t , respectively. We have used the AIC criterion to select the optimal models out of numerous competitors [10]. The optimal model minimizes the AIC.

3.1 Special Case of S5W9x

Using the AIC criterion, the best model has the form:

$$\begin{aligned}
 y_t = \text{const} &+ a_1 y_{t-1} + \dots + a_4 y_{t-4} + a_6 y_{t-6} + a_7 y_{t-7} + b_0 w_t + b_1 w_{t-1} + \dots + b_4 w_{t-4} \\
 &+ c_1 e_{t-1} + c_2 e_{t-2} + d_0 s_t + d_1 s_{t-1} + d_2 s_{t-2} + \text{noise}.
 \end{aligned} \tag{3}$$

The optimal model (3) is referred to as All.step. It points clearly to great dependence of the present eye-gaze on its past. Observe that y_{t-5} is not included in (3).

To have the right perspective, it is helpful to compare the optimal model All.step with M0 in (2), the intuitive model which uses only present arm covariates, and with the model which uses only the past gaze $y_{t-1}, y_{t-2}, \dots, y_{t-12}$, referred to as My1...y12. It is also interesting to consider a model which depends on the immediate past gaze response y_{t-1}, y_{t-2} only. Call this model My1y2.

From Table 2, as compared with M0, the optimal model All.step exhibits a much smaller AIC, much larger R^2 , and much smaller residual standard error. This pattern was observed in all other cases across subjects, movements, and conditions. Model M12 has a substantially larger AIC than All.step, but approximately the same R^2 and residual standard error. Apparently, the inclusion of movement covariates in addition to past gaze covariates brings about further AIC reduction, and helps in fine-tuning the optimal gaze/movement model. Conspicuously, model My1y2, which uses the immediate past response of 16 milliseconds only, accounts for about 97% of the variation in the eye-gaze data. In many cases it is well above 98% This has been observed time and again in many cases.

Table 2: Comparison between All.step, M0, and M12

Model	AIC	R^2	Resid SE
M0	6929.502	0.5164	2.0800
My1y2	2073.291	0.9763	0.4601
My1...y12	2044.084	0.9770	0.4546
All.step	1912.636	0.9789	0.4358

More dramatic is the improvement in the one step ahead prediction shown in Figure 5. Evidently M0 gives a much coarser fit compared with that of All.step.

Next, let us examine the residuals. Figure 6 shows the estimated autocorrelation of the M0 residuals. The autocorrelation points to great stochastic dependence and is completely out of the horizontal bounds, whereas it is completely between the horizontal bounds as expected from white noise (WN). There is a great reduction in autocorrelation going from M0 to All.step, and this was observed throughout.

To reaffirm the whiteness of the All.step residuals, we appeal to the cumulative periodogram (cpgram) in Figure 7. The cpgram of the M0 residuals points to a great emphasis of low frequencies close to the origin. On the other hand, the cpgram of the All.step residuals lies perfectly between the two diagonal confidence bands, as is the case for white noise.

Next we examine the distribution of the All.step residuals in Figure 8. Looking at the figure one is tempted to say that the residuals are normally distributed, but this is not quite the case as the qq-plot shows in Figure 9. This behavior is more in line with that of a mixture of normal distributions shown in Figure 15 for another case.

3.2 Special Case of S3W8x

The previous results where gaze at time t is well explained by a rather long autoregressive component and an additional movement component is typical. Similar findings were observed

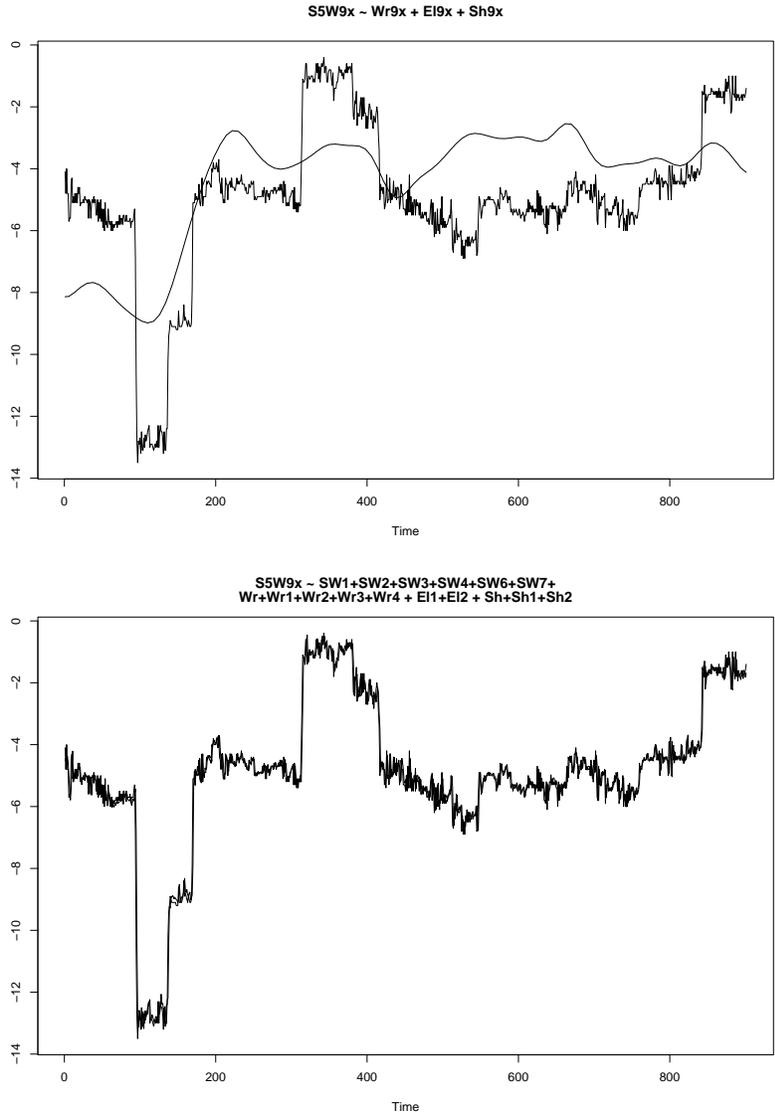


Figure 5: S5W9x: Observed vs fitted values. Top: M0. Bottom: All.step.

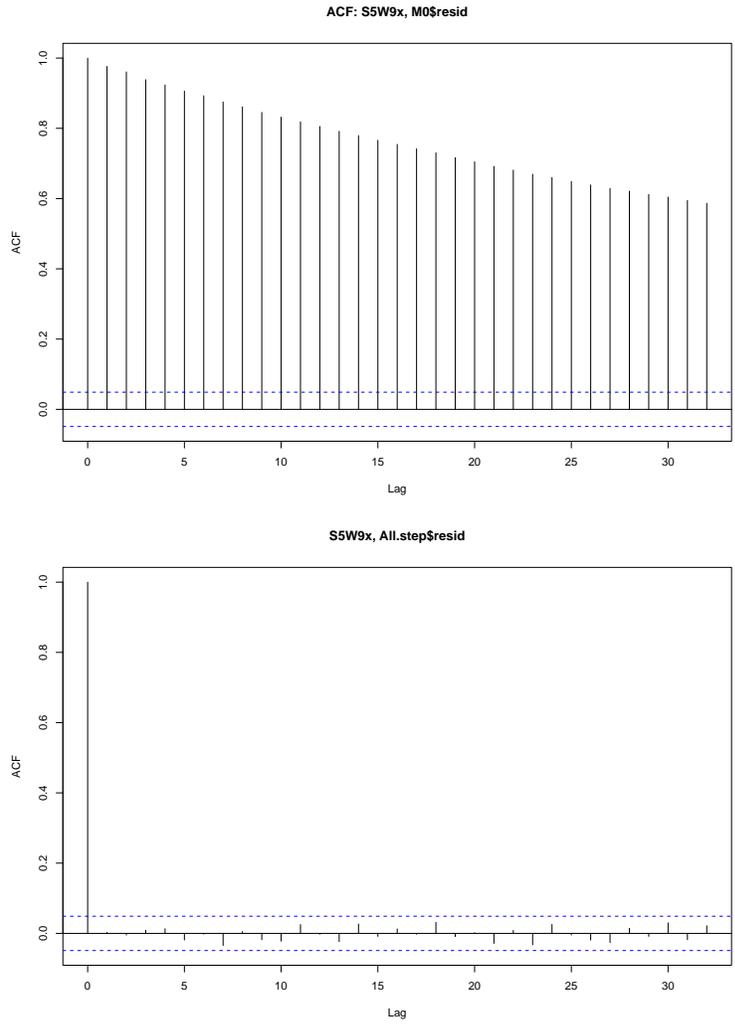


Figure 6: S5W9x. Top: Autocorrelation of the M0 residuals. The autocorrelation points to an inadequate fit. Bottom: Autocorrelation of the All.step residuals. The autocorrelation is within white noise bounds, necessary for an adequate fit.

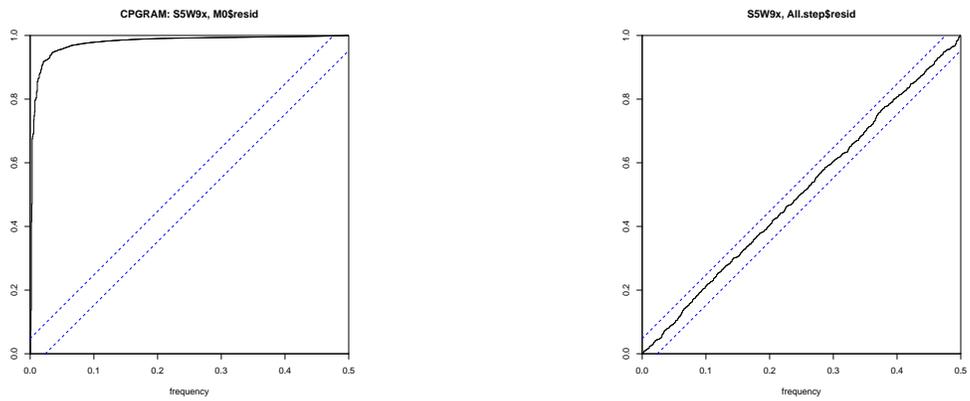


Figure 7: S5W9x. Left: cpgram of the M0 residuals. Right: cpgram of the All.step residuals.

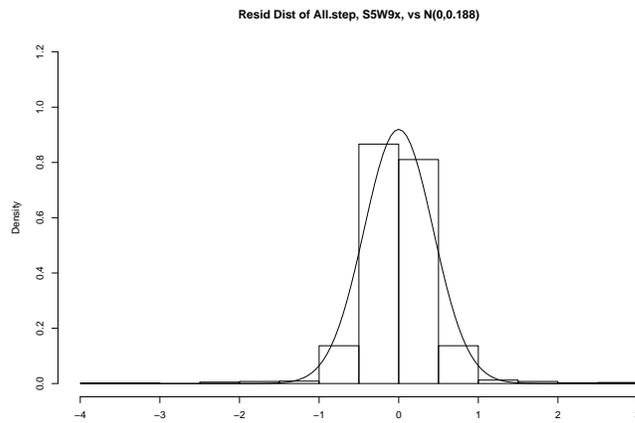


Figure 8: S5W9x: Histogram of the All.step residuals vs the pdf of $N(0,0.188)$.

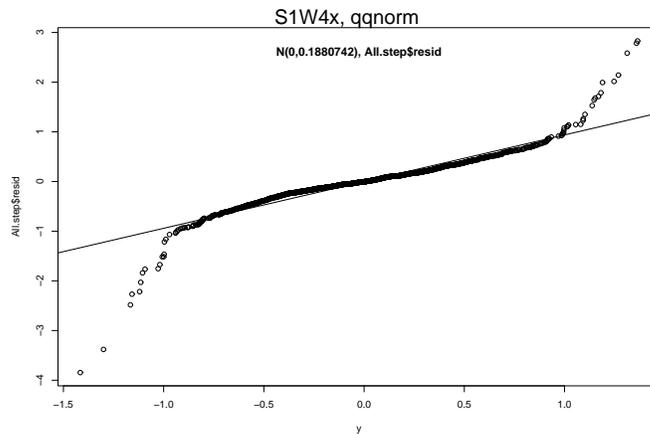


Figure 9: S5W9x: QQ-Norm plot of the All.step residuals.

in other “Watch-x” cases across subjects and movements. In support of this, it is instructive to consider yet another special case, that of S3W8X. In addition we illustrate the effect of removal of certain covariates from an optimal model.

With the same notation as above, the best model which minimizes the AIC is again referred to as All.step and is given as

$$\begin{aligned}
y_t = \text{const} &+ a_1y_{t-1} + \cdots + a_5y_{t-5} + a_8y_{t-8} + a_9y_{t-9} + a_{10}y_{t-10} \\
&+ b_0w_t + b_1w_{t-1} + b_3w_{t-3} + b_4w_{t-4} \\
&+ c_1e_{t-1} + c_2e_{t-2} + \text{noise}
\end{aligned} \tag{4}$$

Notice that the autoregressive part has a “hole” (y_{t-6}, y_{t-7} not included), as does the wrist part (w_{t-3} not included), and that the shoulder did not enter at all. This pattern has been observed in many cases. Still, there are cases where the shoulder was included.

It is interesting to see what happens when certain covariates are removed from the optimal model All.step. That is the effect on the residual cumulative periodogram (cpgram), AIC, R^2 , residual standard error (RSE), and the effect on the plot of observed vs fitted values. In this respect, consider the following models. The model

$$\begin{aligned}
y_t = \text{const} &+ a_1y_{t-1} + \cdots + a_5y_{t-5} + a_8y_{t-8} + a_9y_{t-9} + a_{10}y_{t-10} \\
&+ \text{noise}
\end{aligned} \tag{5}$$

retains only the autoregressive component of All.step, and is called “AR Only”. Next we remove from All.step the autoregressive component to come up with model “WE”, consisting of wrist and elbow components only,

$$y_t = \text{const} + b_0w_t + b_1w_{t-1} + b_3w_{t-3} + b_4w_{t-4} + c_1e_{t-1} + c_2e_{t-2} + \text{noise} \tag{6}$$

Model AR12.WE adds the autoregressive terms y_{t-1}, y_{t-2} to model WE,

$$\begin{aligned}
y_t = \text{const} &+ a_1y_{t-1} + a_2y_{t-2} \\
&+ b_0w_t + b_1w_{t-1} + b_3w_{t-3} + b_4w_{t-4} \\
&+ c_1e_{t-1} + c_2e_{t-2} + \text{noise}
\end{aligned} \tag{7}$$

Model M0 is the same as (2). Evidently, from Table 3, the optimal model requires both a long memory gaze autoregressive part as well as some arm covariates, where the autoregressive part plays a more important role.

The following figures illustrate the improvement achieved by adding the autoregressive gaze component to the arm component. Figures 10 and 11 display the residual autocorrelation reduction obtained by going from model WE to model All.step, and Figures 12 and 13 show the corresponding improvement in goodness of fit. Model WE traces the general smooth version of the gaze data, whereas model All.step traces the data in detail.

Figure 14 shows the great symmetry of the residuals distribution. The distributions of the residuals from many other Watch-x cases appear equally symmetric. The qq-plot in Figure 15 suggests that the residual distribution may be approximated by a mixture of two normal distributions, $N(0,0.16)$ and $N(0,2.25)$, with weights approximated by 0.98 and 0.02, respectively.

Table 3: Comparison between All.step and competitors.

Model	Resid CPGRAM	AIC	R^2	Resid SE
WE	All out	2769.9	0.8989	1.3560
M0	All out	2738.7	0.9020	1.3320
AR Only	All in	1207.1	0.9857	0.5106
AR12.WE	Most in	1087.1	0.9877	0.4737
All.step	All in	1013.6	0.9889	0.4508

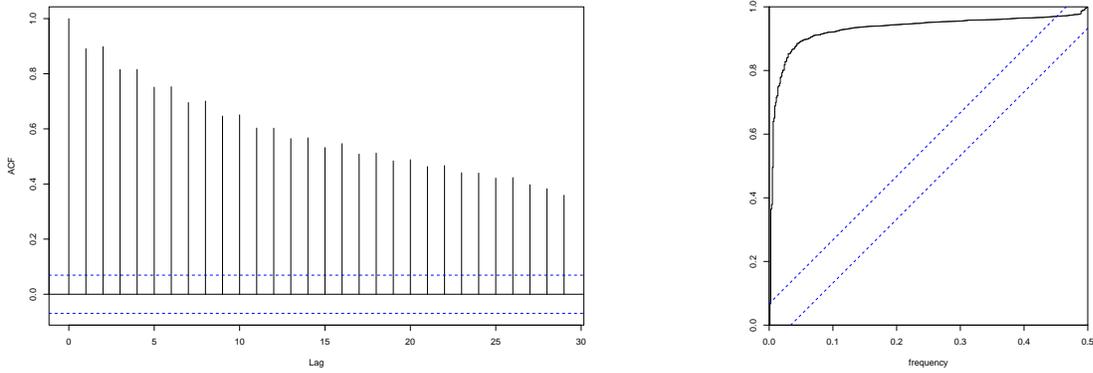


Figure 10: S3W8x: Residual acf (left), and cpgram (right) of model WE.

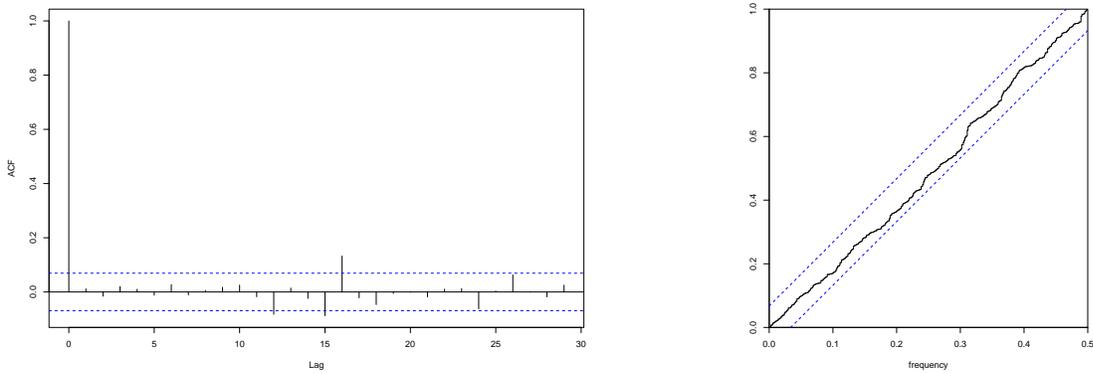


Figure 11: S3W8x: Residual acf (left), and cpgram (right) of model All.step

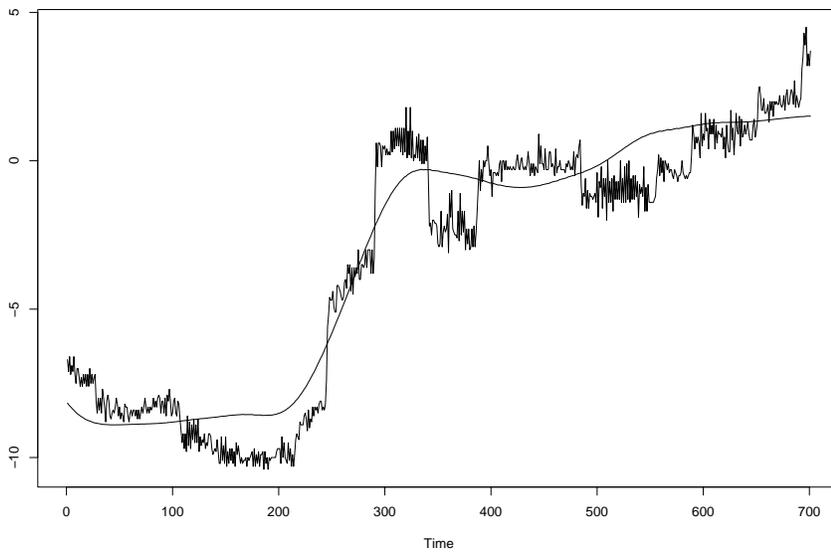


Figure 12: S3W8x: Observed vs fitted values for model WE.

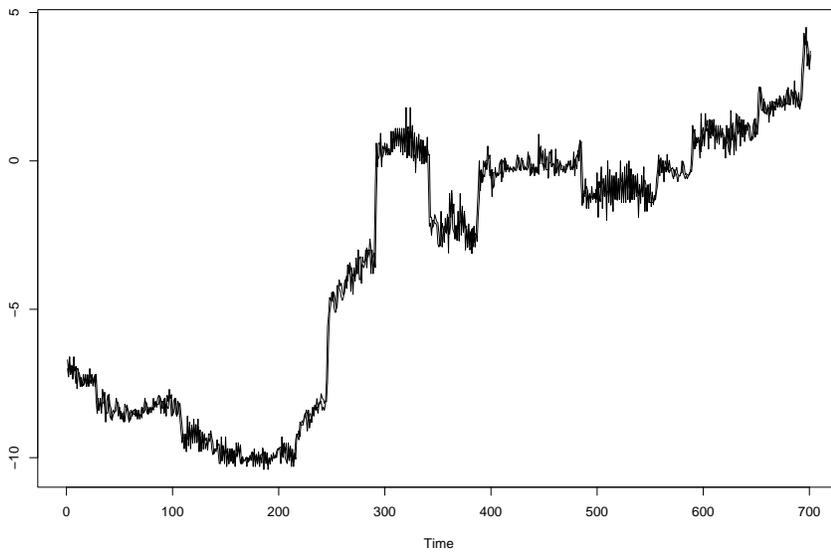


Figure 13: All.step: Observed vs fitted values.

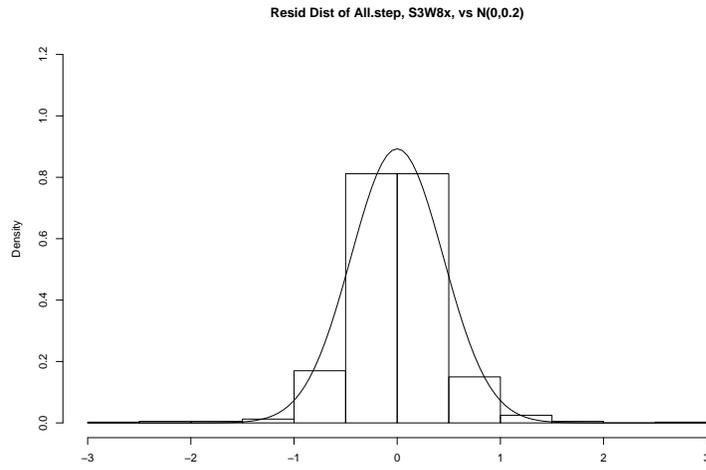


Figure 14: S3W8x: Residual histogram of model All.step vs the pdf of $N(0,0.2)$

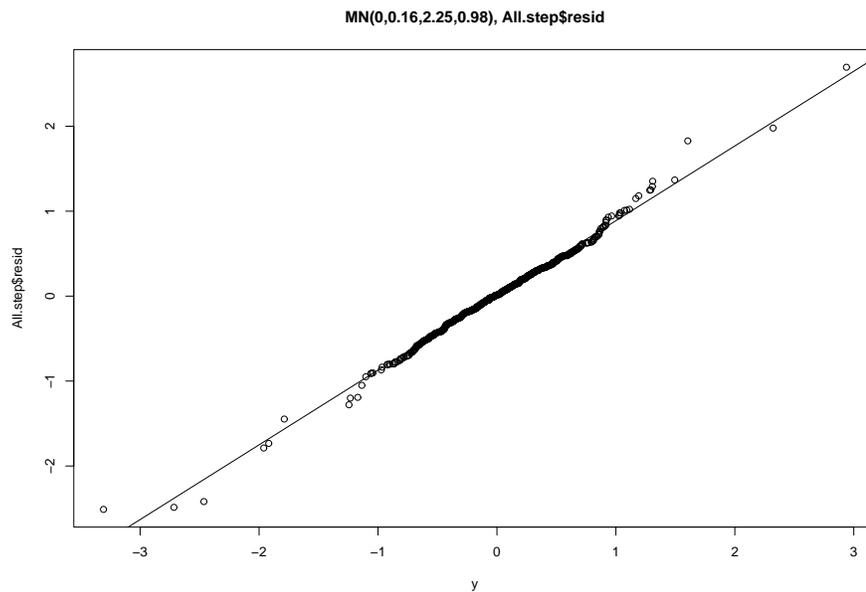


Figure 15: S3W8x: QQ-Plot of the All.step residuals and data from a normal mixture.

3.3 Special Case of S3W6y

The results for the y-coordinate of the gaze are observed to be in line with that of the x-coordinate of the gaze. The best model based on the smallest AIC value has a long autoregressive part as well as a movement component, a universal phenomenon across all subjects and movements. The special case of S3W6y is one such example. The best model All.step is given by

$$y_t = const + a_1y_{t-1} + \dots + a_7y_{t-7} + a_9y_{t-9} + a_{12}y_{t-12} + d_2s_{t-2} + noise \quad (8)$$

As before the autoregressive part has a hole as $y_{t-8}, y_{t-10}, y_{t-11}$ are not included. It is also interesting to note that shoulder is the only covariate that enters the model. In most cases wrist and elbow take precedence over the shoulder component.

Comparing the different models with the model All.step gives similar results as before. There is a great reduction in the AIC when an autoregressive part is added to the model. Also observe that the addition of the shoulder component helps increase the R^2 and decrease the residual SE by a small margin.

Table 4: Comparison between All.step and some other models.

Model	AIC	R^2	Resid SE
M0	5500.690	0.5307	3.192
M12	1993.281	0.9829	0.6125
AR Only	1989.407	0.9828	0.6123
All.step	1978.648	0.9930	0.6089

The following figures compare the performance of the models M0 and All.step. Figures 16 and 17 show that residuals from model All.step behave like white noise as compared to model M0 residuals which are highly correlated. Figure 18 shows that the fitted time series traces the observed time series very closely.

3.4 Special Case of S2W4r

The distance coordinate is the square root of the sum of squares of the x and y- coordinates, $r = \sqrt{x^2 + y^2}$. Experience indicates the results are no different from before.

As an example, consider the case of S2W4r. The distance r-coordinate, denoted by y_t , was regressed on past r-coordinates and movement covariates as before. The best model, All.step, with the least AIC is given by

$$\begin{aligned} y_t = const + & a_1y_{t-1} + \dots + a_7y_{t-7} + a_{10}y_{t-10} + a_{11}y_{t-11} + a_{12}y_{t-12} \\ & + b_0w_t + b_1w_{t-1} + \dots + b_4w_{t-4} \\ & + c_0e_t + c_1e_{t-1} + noise \end{aligned} \quad (9)$$

The All.step model shows a hole (y_{t-8}, y_{t-9} not included). The wrist component has all its covariates and elbow has a hole(e_{t-2}) as well. The shoulder component is absent like in many

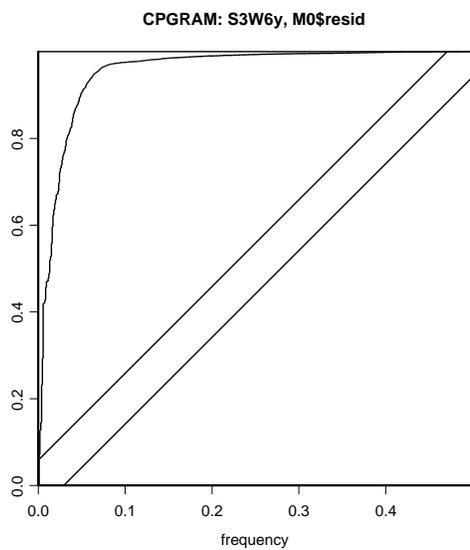
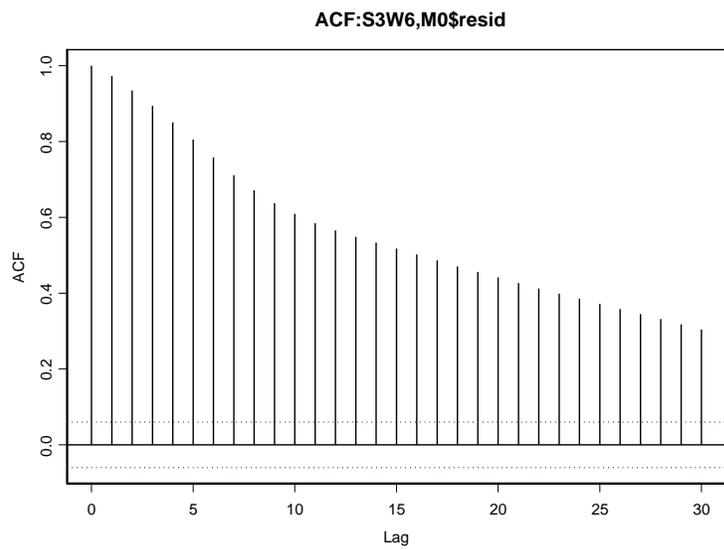


Figure 16: S3W6y: Residual acf (top), and cpgram (bottom) of model M0.

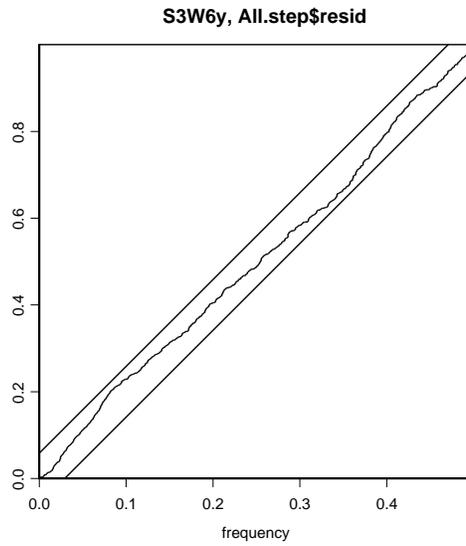
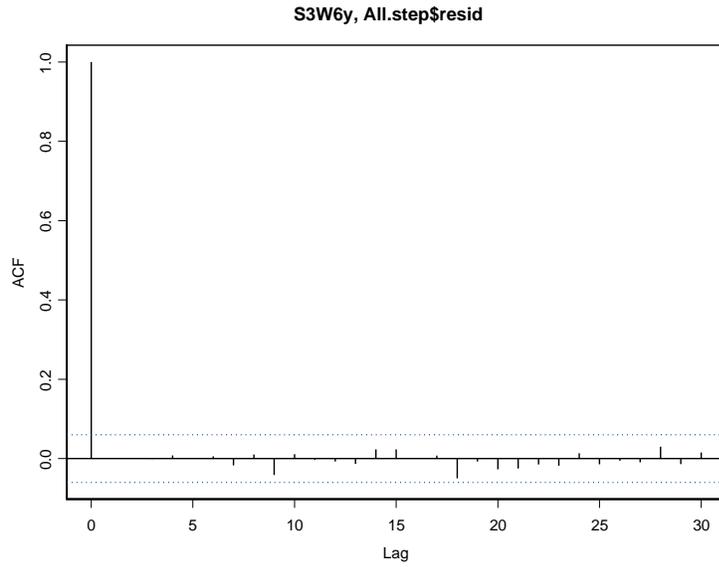


Figure 17: S3W6y: Residual acf (top), and cpgram (bottom) of model All.step

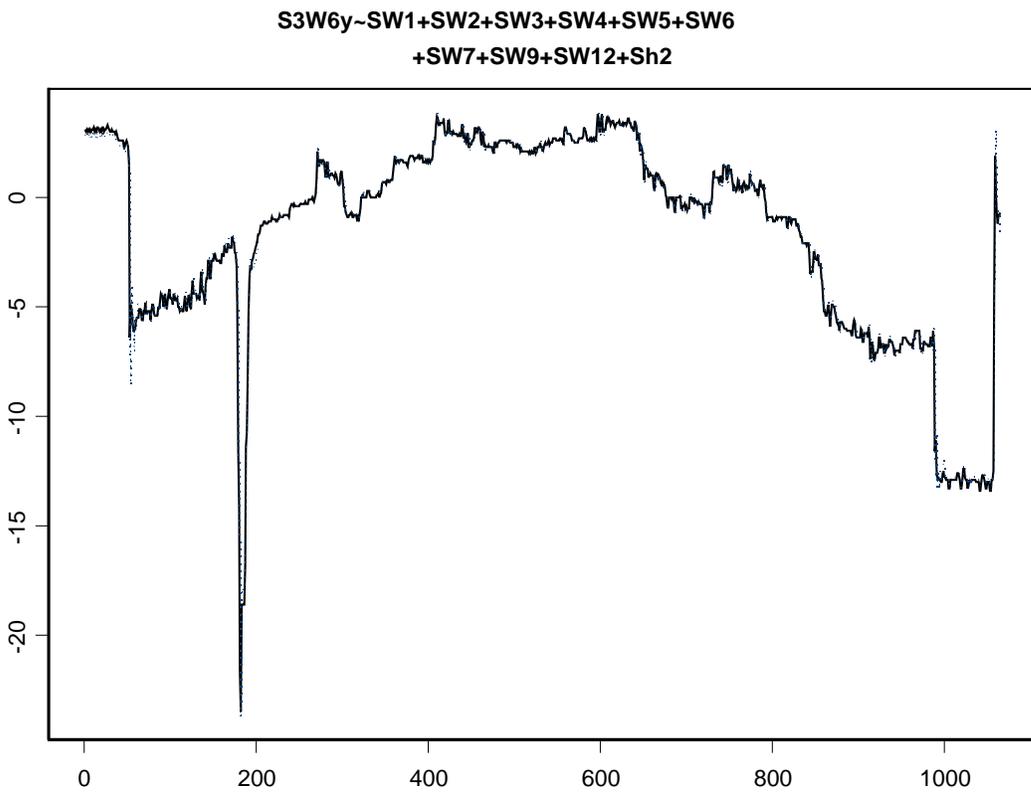


Figure 18: S3W6y: Observed vs fitted values for model All.step.

other cases. As before, the optimal model includes the first two components y_{t-1}, y_{t-2} , as do all the other optimal models corresponding to the different gaze-movement cases. It is interesting to see what happens when y_{t-1}, y_{t-2} are removed from All.step. So, consider the model

$$\begin{aligned}
y_t = \text{const} &+ a_3y_{t-3} + \cdots + a_7y_{t-7} + a_{10}y_{t-10} + a_{11}y_{t-11} + a_{12}y_{t-12} \\
&+ b_0w_t + b_1w_{t-1} + \cdots + b_4w_{t-4} \\
&+ c_0e_t + c_1e_{t-1} + \text{noise}
\end{aligned} \tag{10}$$

and refer to it as All.step-12.

Table 5 summarizes the results of several models. Model All.step-12 shows a high AIC, relatively low R^2 , and high residual standard error, but a much higher AIC value as compared to the optimal model All.step.

Table 5: Comparison between All.step and competitors.

Model	AIC	R^2	Resid SE
WE	5678.775	0.2095	2.484
M0	5896.906	0.0480	2.722
AR Only	1428.463	0.9761	0.4328
All.step-12	3171.945	0.8842	0.9005
All.step	1419.040	0.9765	0.4299

The analysis of residuals for the two models All.step and All.step-12 is illustrated in the plots. Figures 19 and 20 shows there is considerable reduction in correlation of the residuals. Notice that the cumulative periodogram plot of All.step-12 shoots way above the upper bound in Figure 21. This is rectified in Figure 22 by the inclusion of y_{t-1}, y_{t-2} . This exact same behavior has been observed time and again in many other cases which shows that y_{t-1}, y_{t-2} are essential for an adequate fit.

ACF: S2W4r, All.step-12\$resid

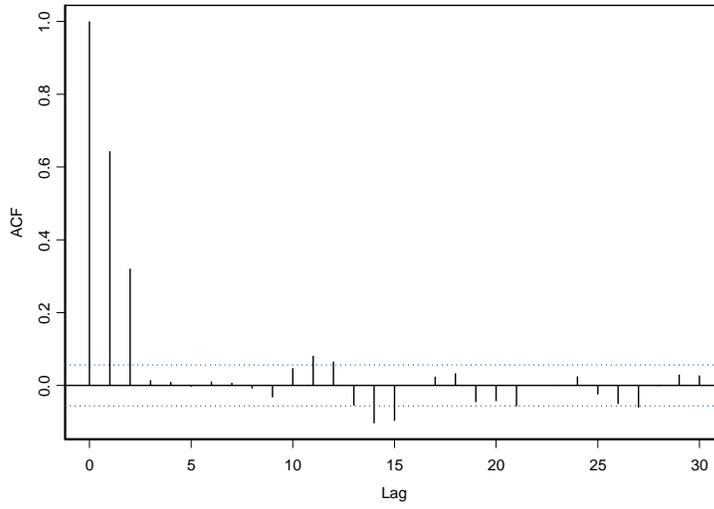


Figure 19: S2W4r: Residual acf of model All.step-12.

ACF: S2W4r, All.step\$resid

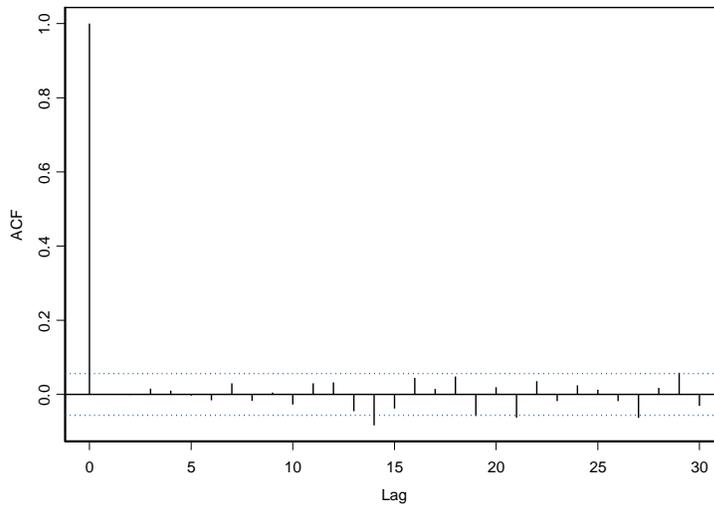


Figure 20: S2W4r: Residual acf of model All.step.

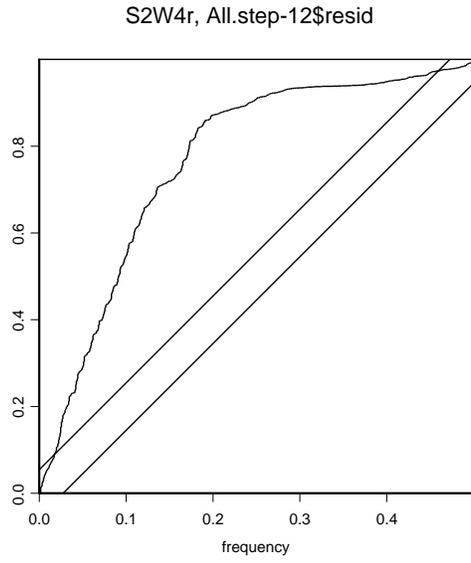


Figure 21: S2W4r: Residual acf of model All.step-12.

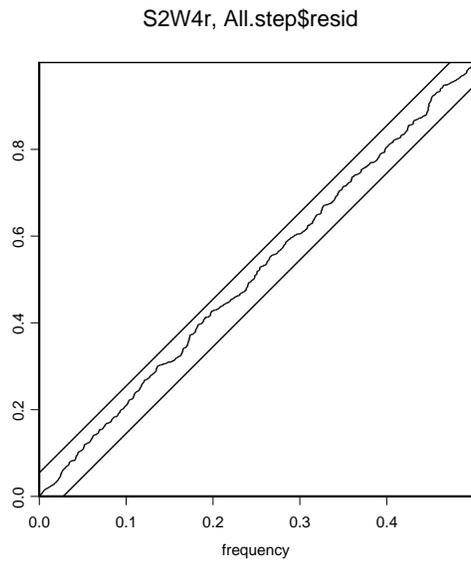


Figure 22: S2W4r: Residual acf of model All.step.

3.5 Special Case of S2W4r: New using all 1241 obs.

Our experience indicates that the results from the distance r-coordinate, $r = \sqrt{x^2 + y^2}$, are similar to those from the x- and y- coordinates. A typical example of r-coordinate eye-gaze is given in terms of Subject 2, Watch, movements 4 (S2W4r). The eye-gaze time series is shown in Figure 23 together with the wrist, elbow, and shoulder arm inputs.

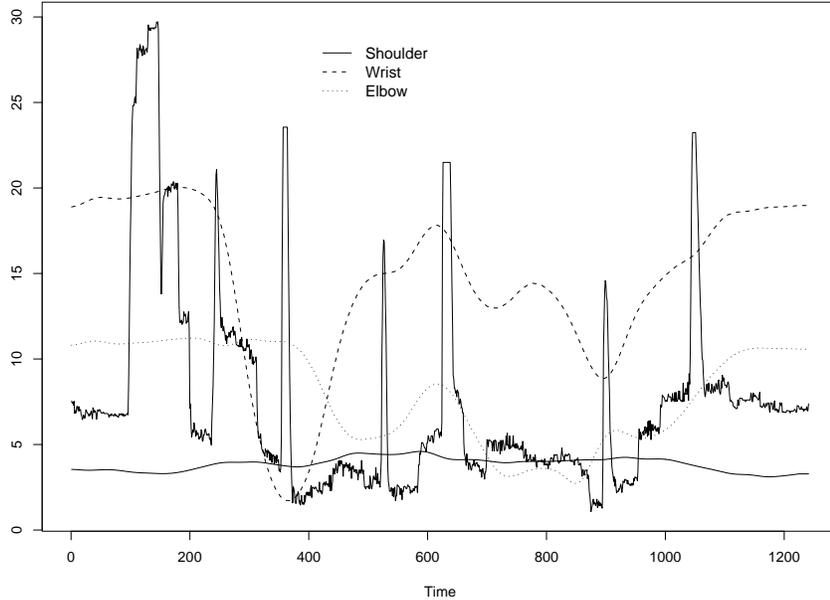


Figure 23: A time series from subject 2, movement 4, r-coordinate, under “Watch”, and arm inputs sampled at 120Hz.

The r-coordinate eye-gaze, again denoted by y_t , was regressed on past r-coordinates from eye-gaze and from movement covariates as before. The optimal model, All.step, with the smallest AIC, is given by

$$y_t = const + a_1 y_{t-1} + \dots + a_7 y_{t-7} + a_{11} y_{t-11} + b_2 w_{t-2} + c_1 e_{t-1} + noise \quad (11)$$

The optimal All.step model shows a hole as $y_{t-8}, y_{t-9}, y_{t-10}$ are not included. Wrist and elbow components are present but shoulder components are absent like in many other cases.

3.5.1 The Effect of Removing y_{t-1}, y_{t-2}

The optimal model includes the first two components y_{t-1}, y_{t-2} , as do all the other optimal models corresponding to the different gaze-movement cases, regardless of the coordinate type. Recall that at 120Hz this refers to the immediate past 16 milliseconds. It is interesting to see what happens when y_{t-1}, y_{t-2} are removed from the optimal All.step. Thus, we consider

the model

$$y_t = const + a_3y_{t-3} + \dots + a_7y_{t-7} + a_{11}y_{t-11} + b_2w_{t-2}c_1e_{t-1} + noise \quad (12)$$

and refer to it as All.step-y1y2.

Table 6 summarizes the results of several models. As in all cases, M0 is clearly far from being optimal. Model All.step-y1y2 has a relatively low R^2 , relatively high residual standard error, but a much higher AIC value as compared to My1y2, let alone the optimal model All.step. Evidently, the removal of the immediate past 16 milliseconds from the optimal model greatly diminishes the quality of the model. This has been observed throughout. Further evidence of this deterioration is seen from a comparison of the residuals in Figures

Table 6: Comparison between All.step and competitors.

Model	AIC	R^2	Resid SE
M0	7609.221	0.2735	5.1780
All.step-y1y2	4972.646	0.9139	1.7860
My1y2	2667.840	0.9864	0.7075
My1...y12	2647.410	0.9869	0.6989
All.step	2637.966	0.9869	0.6968

24-27. Interestingly, in terms of AIC, R^2 , and residual standard error, My1y2 with only two autoregressive components y_{t-1}, y_{t-2} is very close to My1...y12 with twelve autoregressive components y_{t-1}, \dots, y_{t-12} . This gives further evidence of the significance of the immediate past as embodied in y_{t-1}, y_{t-2} . We have seen this in the earlier case of S5W9x, and in many other cases as well.

Figures 24 and 25 show there is considerable reduction in the autocorrelation of the residuals. In addition, the cumulative periodogram plot of All.step-y1y2 shoots way above the upper bound in Figure 26. This is rectified in Figure 27 by the inclusion of y_{t-1}, y_{t-2} . This exact same behavior has been observed time and again in all other cases, which shows that y_{t-1}, y_{t-2} are essential for an adequate fit. This may be interpreted to mean that memory of the immediate past on the order of 16 milliseconds is essential for movement prediction.

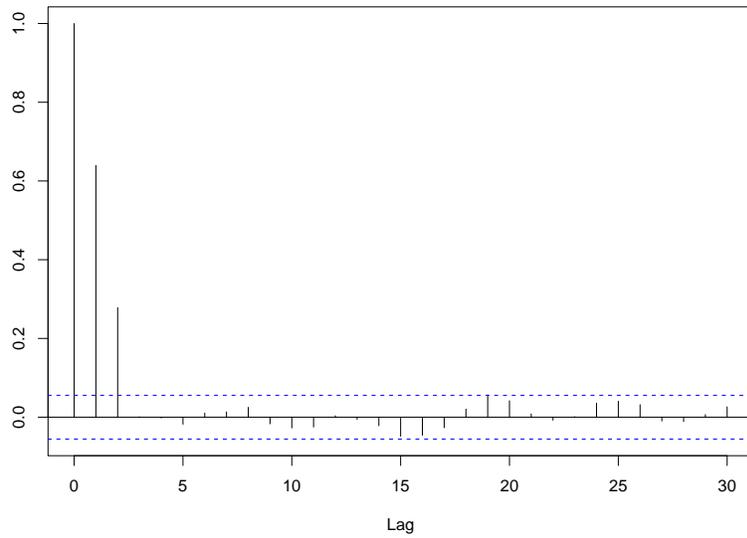


Figure 24: S2W4r: Residual autocorrelation of model All.step-y1y2. The autocorrelation points to an inadequate fit.

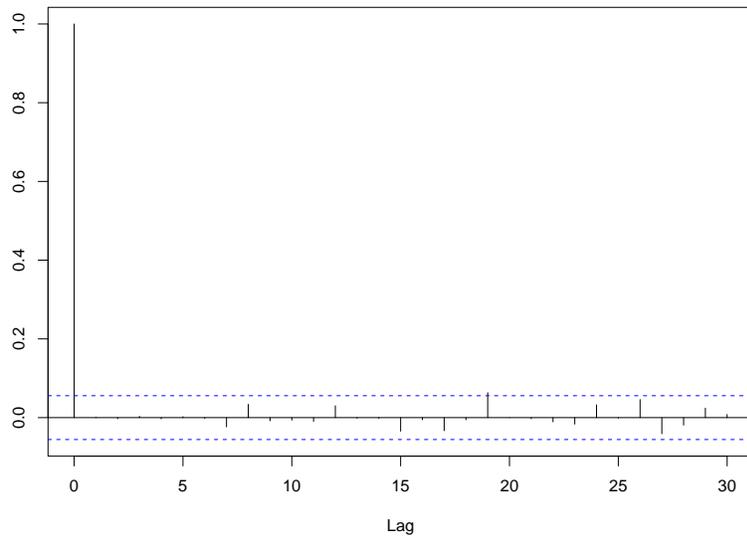


Figure 25: S2W4r: Residual autocorrelation of the optimal model All.step. The autocorrelation resembles that of white noise, necessary for an adequate fit.

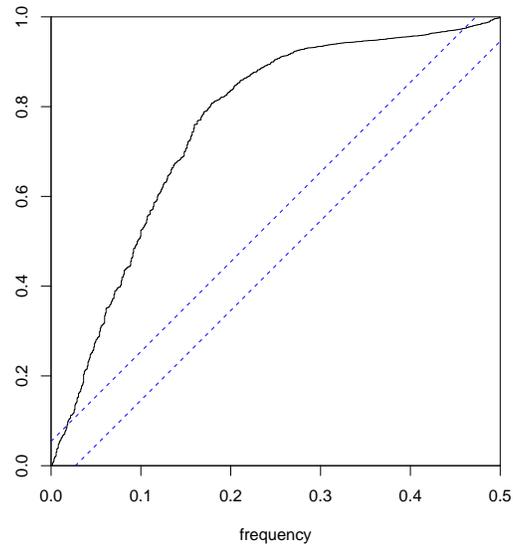


Figure 26: S2W4r: Residual cumulative periodogram of model All.step-y1y2. The cumulative periodogram points to correlated residuals.

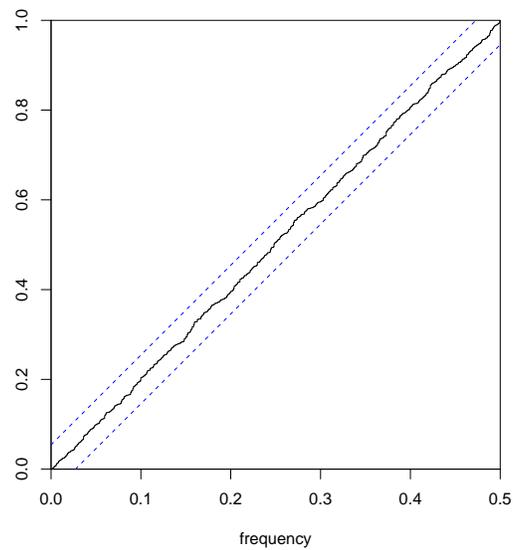


Figure 27: S2W4r: Residual cumulative periodogram of the optimal model All.step within white noise bounds, necessary for an adequate fit.

3.6 Special Case of S4I7x

The best model All.step is given by

$$\begin{aligned}
 y_t = \text{const} &+ a_1y_{t-1} + a_2y_{t-2} + a_3y_{t-3} + a_5y_{t-5} + a_6y_{t-6} + a_9y_{t-9} + a_{10}y_{t-10} + a_{11}y_{t-11} \\
 &+ b_0w_t + b_1w_{t-1} + b_3w_{t-3} \\
 &+ d_0sh_t + d_1sh_{t-1} + \text{noise}
 \end{aligned}
 \tag{13}$$

Note that the elbow component is absent from the optimal model.

The previous sections point out the importance for the first two components of the autoregressive part of the optimal model. It has been noticed that the cumulative periodogram (cpgram) shoots up above the upper bound in their absence. Consider now the comparison of the All.step with the model “AR12 Only” given by

$$y_t = \text{const} + a_1y_{t-1} + a_2y_{t-2} + \text{noise}
 \tag{14}$$

A comparison of various models is given in Table 7 below.

Table 7: Comparison between All.step and competitors.

Model	AIC	R^2	Resid SE
M0	4123.587	0.8005	1.518
WE	4088.025	0.8074	1.493
AR Only	1978.404	0.9710	0.581
AR12 Only	2057.813	0.9683	0.6045
All.step	1946.82	0.9719	0.5725

From the table, the R^2 value for the AR12 Only model suggests that y_{t-1}, y_{t-2} account for 96.8% of the variability in the data compared to a 97.2% in the case of the best model. More important is the large drop in the corresponding AIC values.

The following figures compare the performances of model AR12 Only and All.step. The autocorrelation and cumulative periodogram figures 28 and 29 show the improvement of the model AR12 Only when the other AR components and arm components are added to it. Figures 30 and 31 however are difficult to differentiate in this case.

4 Removal of y_{t-1}, y_{t-2}

Appear in all models!!! y_{t-1}, y_{t-2} not far from optimal in the sense it explains about 97% – 98% of the total variation.

As was noted in the case of S2W4r, the removal of y_{t-1}, y_{t-2} is detrimental to the adequacy of the model. We see this again from Figure 32 where the cumulative periodograms get off bound displaying a surprisingly homogeneous “badness”. Since the data were sampled at 120Hz, this may be interpreted to mean that the previous 16-20 milliseconds are critical for prediction of eye-gaze in response to arm movement. In other words, in general, our models predict on a scale of about 1/10 of a second, but for this to be successful, the previous immediate eye-gaze on a scale of about 16-20 milliseconds must be part of the model.

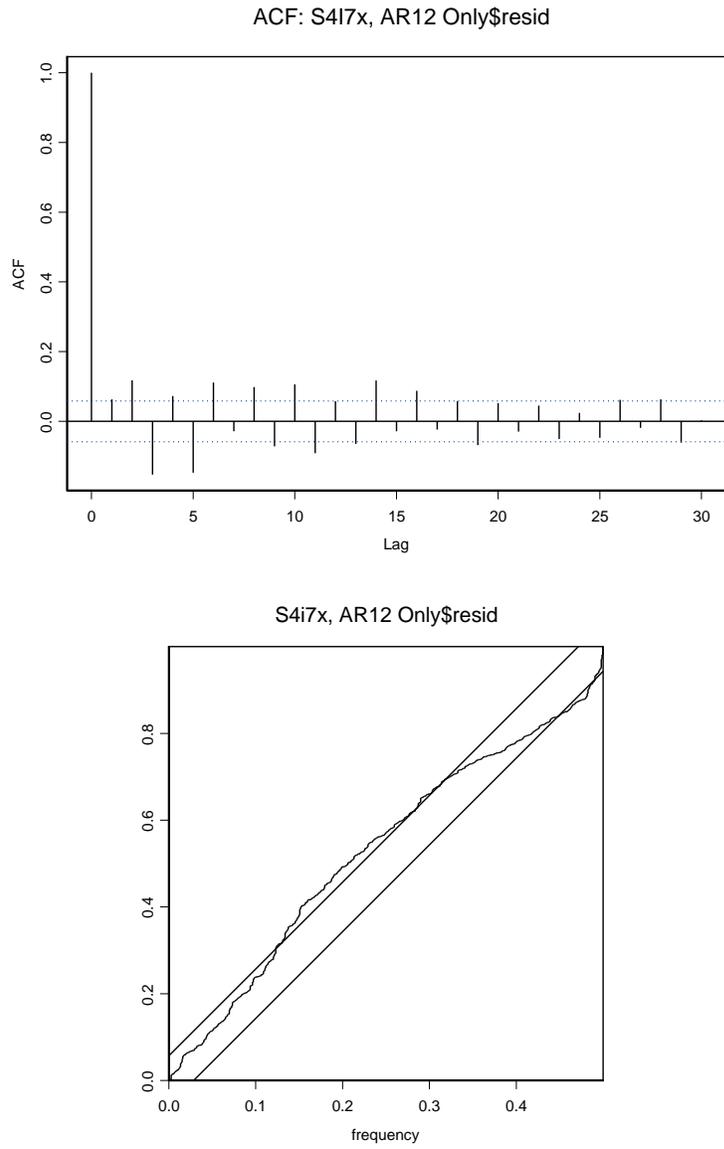


Figure 28: S4I7x: Residual acf (top), and cpgram (bottom) of model AR12 Only.

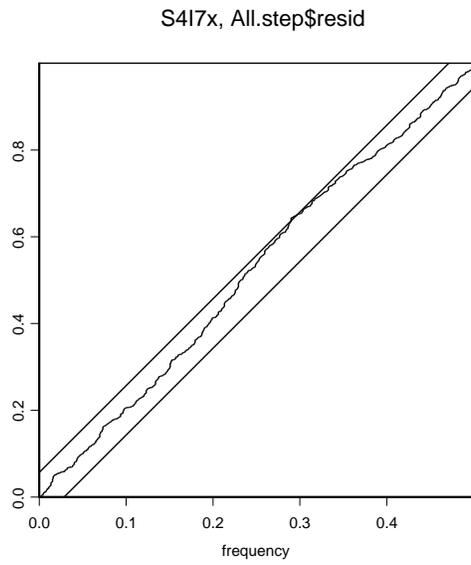
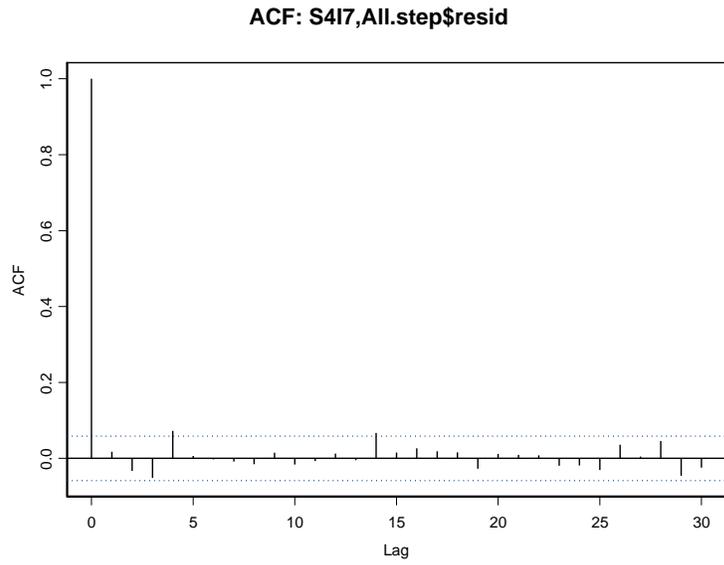


Figure 29: S4I7x: Residual acf (top), and cpgram (bottom) of model All.step

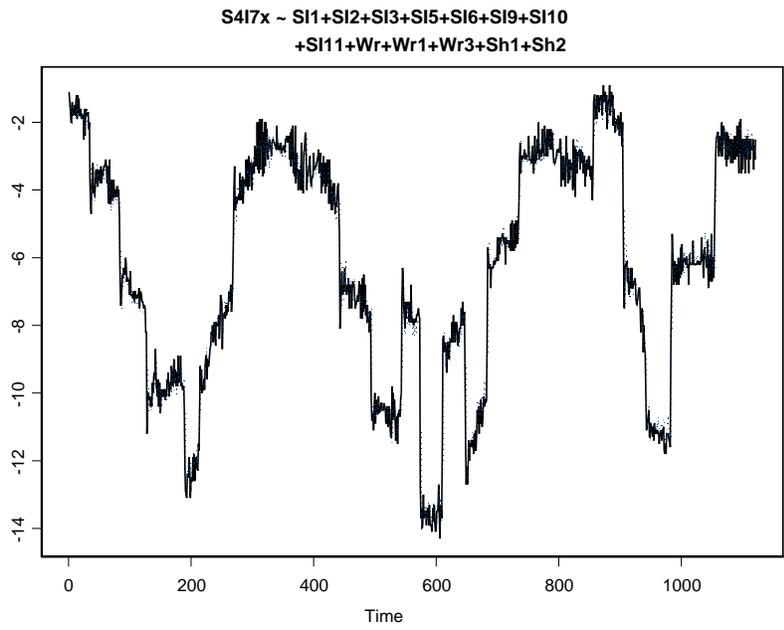


Figure 30: S4I7x: Observed vs fitted values for model All.step.

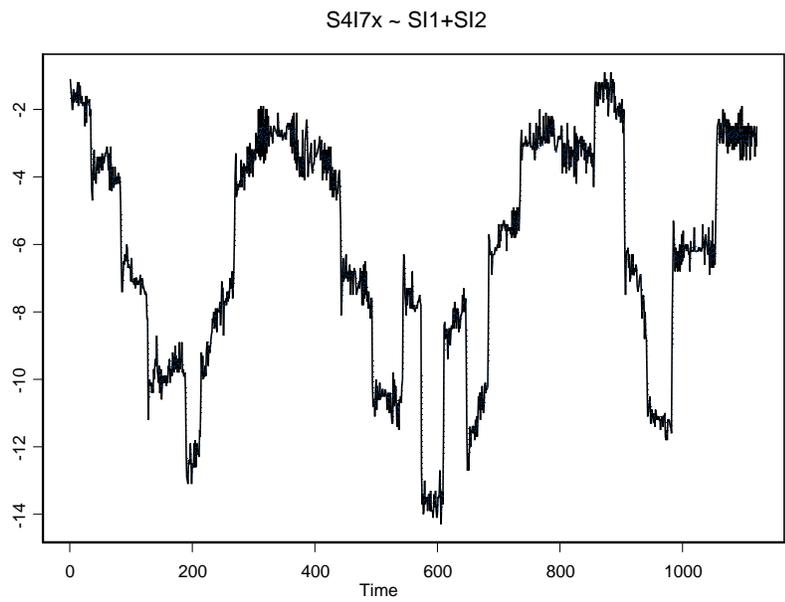


Figure 31: S4I7x: Observed vs fitted values for model AR12 Only.

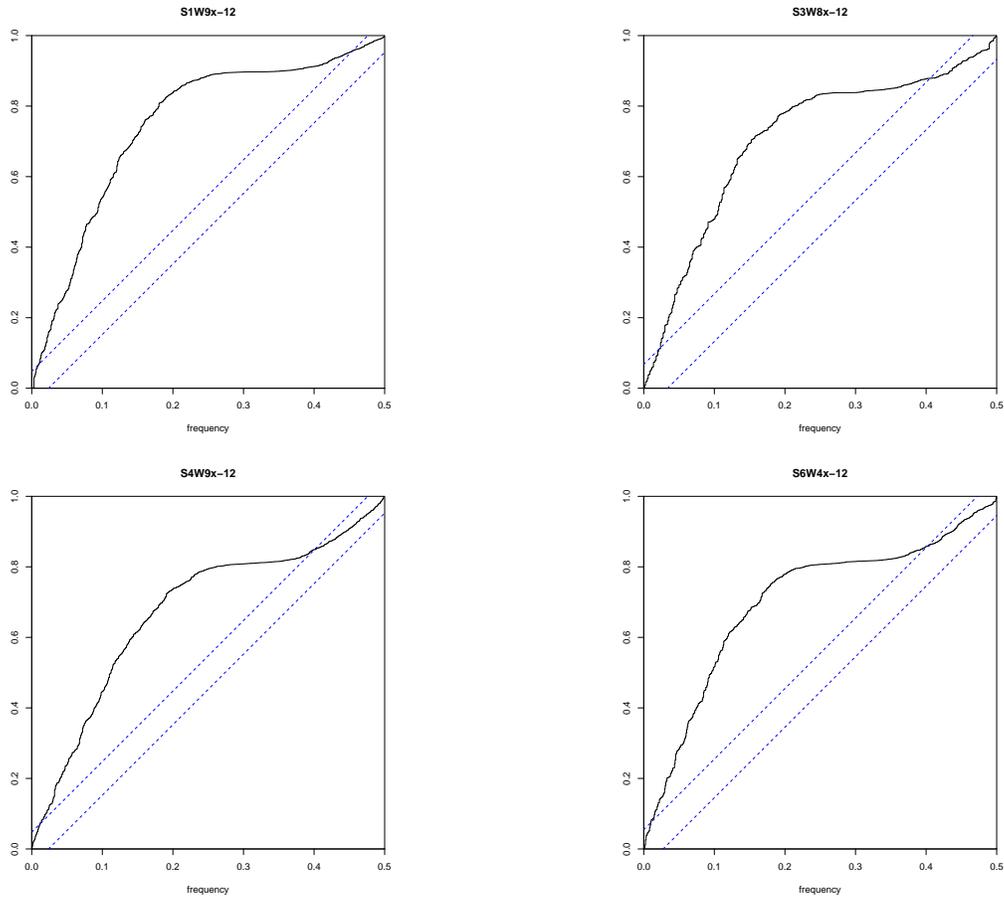


Figure 32: Cograms of All.step-12 cases: S1W9x, S3W8x, S4W9x, S6W4x. The omission of y_{t-1}, y_{t-2} causes the cogram to get out of bound.

4.1 Summary of the Regression Results

To predict need past!!! Explain

In relating eye-gaze to arm movement, the picture which emerges from the previous cases, and from the additional analysis of many other cases regarding both Watch and Imitate shows that:

- a. The optimal model in each case has a long autoregressive part, plus “some help” mainly from the wrist, less from the elbow data, and even less from the shoulder data. The autoregressive component points to a long memory inherent in the gaze time series.
- b. The analysis shows that when y_{t-1}, y_{t-2} are not included in a model, the cpgram shoots way above the upper bound. Therefore, these two components are essential for model adequacy and well behaved residuals.
- c. The unusually high R^2 in All.step on the order of 98% and more means that the model explain at least 98% of the variability in the data. This is due mainly to the autoregressive component.
- d. The cpgram from the AIC-optimal All.step is well within the white noise bounds as seen for example in Figure 7. There were cases where the autoregressive part by itself was insufficient for achieving this ideal, and the inclusion of arm movement components was necessary.
- e. There is a great improvement in model fit going from a model containing arm movement only, M0, to the optimal model All.step containing a long memory autoregressive part.

References

- [1] Box, G.E.P, Jenkins, G.M., Reinsel, G.C. (1994). *Time series analysis and Forecasting*, 2nd Ed. Prentice hall, Englewood Cliffs, NJ.
- [2] Barkoulas, J., Baum, C. and Oguz, G. (1997). Stochastic long memory in traded goods prices. *Boston college working papers in economics*, 1997, 349.
- [3] Burlando1, P., Montanari1, A. and Rosso1, R. (1996). Modelling hydrological data with and without long memory, *Meccania*, 31(1), 87-101.
- [4] Chen, Y., Ding, M. and Kelso, J. A. Scott. (1997). Long Memory Processes ($1/f^\alpha$ Type) in Human Coordination, *Physical Review Letters*, 79(22), 4501–4504.
- [5] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes, *Annals of statistics*, 17, 1749-1766.
- [6] Fokianos, K., Kedem, B., Qin, J., and Short, D. A. (2001). A semiparametric approach to the one-way layout. *Technometrics*, 43, 56-65.
- [7] Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, *Annals of statistics*, 14, 517-532.
- [8] Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle’s estimate, *Probability theory and related fields*, 86, 87-104.

- [9] Galbraith, J. and Zinde-Walsh, V. (2001). Autoregression-based estimators for ARFIMA models, CIRANO Working Papers, 2001s-11.
- [10] Kedem, B. and Fokianos, K. (2002). *Regression Models for Time Series Analysis*, Wiley, New York.
- [11] Semi-parametric cluster detection. *J. Stat. Theory and Practice*, 1, 49-72.
- [12] Lu, G. (2007). Asymptotic Theory for Multiple-Sample Semiparametric Density Ratio Models and its Application to Mortality Forecasting. Ph.D. Dissertation, University of Maryland, College Park.
- [13] McCulloch, C.E. and Searle, S.R. (2001). *Generalized, Linear, and Mixed Models*, Wiley, New York.
- [14] Noy, L.(2009). From Action to Perception and Back Again: Studying the features of movement imitation. PhD Dissertation, Weizmann Institute of Science, Rehovot, Israel.
- [15] Palma, W. (2007). *Long-memory time series, Theory and methods*, John Wiley and sons, USA.
- [16] Qin, J., and Zhang, B. (1997). A goodness of fit test for logistic regression models based on case-control data. *Biometrika*, 84, 609-618.
- [17] Sibbertsen, P. (2004). Long memory in volatilities of German stock returns, *Empirical Economics*, 29, 477-488.
- [18] Whittle, P. (1953). Estimation and Information in Stationary Time Series, *Arkiv for Matematik*, 2, 423-434.
- [19] Zhang, B. (2000c). A goodness of fit test for multiplicative-intercept risk models based on case-control data. *Statistica Sinica*, 10, 839-865.