





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

(1)

Estimating Distributed Lag Coefficients when there  
are Errors in the Observed Time Series

Melvin J. Hinich\*

Abstract

Suppose that two stationary time series satisfy the linear relationship  $y(n) = \sum_{m=0}^{\infty} h(m)x((n-m)\tau)$ . Estimating the distributed lag coefficients  $\{h(m)\}$  from a sample of the two processes when  $\{x(n\tau)\}$  and  $\{y(n\tau)\}$  are measured with error is a statistical problem that is frequently encountered in physical science, engineering, and social science applications. In the engineering and science literature the distributed lags are called the impulse response weights of a causal linear filter. A least squares fit of the model gives biased estimates of the coefficients for this time series version of the errors-in-variables problem. This paper presents approximately unbiased estimators of a scalar multiple of the coefficients. The large sample variances of these estimators are identical, and are of the order  $O(N^{-1})$  where  $N$  is the sample size. The accuracy and practicality of the estimating procedure is illustrated by some simulation results.

\*This work was supported by the Office of Naval Research (Statistics and Probability Program) under contract. I have benefitted from many useful discussions of this subject with Warren Weber, and I am grateful to James McGlone for his computational assistance.

08 01 17 180

ADA 123460

DTIC FILE COPY

Estimating Distributed Lag Coefficients when there  
are Errors in the Observed Time Series

Melvin J. Hinich  
Virginia Polytechnic Institute & State University  
Blacksburg, Virginia

Introduction

Suppose that two discrete-time stationary random processes satisfy the linear relationship  $y(t_n) = \sum_{m=-\infty}^{\infty} h(t_m)x(t_n - t_m)$  where  $t_n = n\tau$ . Estimating the sequence  $\{h(t_m)\}$  from a sample of the two processes is a statistical problem that is frequently encountered in physical science, engineering, and social science applications. Often  $\{x(t_n)\}$  and  $\{y(t_n)\}$  are measured with error, and are of the form  $x(t_n) = x(t_n) + u(t_n)$  and  $y(t_n) = y(t_n) + \varepsilon(t_n)$  where  $\{\varepsilon(t_n)\}$  and  $\{u(t_n)\}$  are noise processes. Then a least squares fit of the model yields biased estimates of the  $h(t_n)$ . These biases are examples of the errors-in-variables biasing problem for a linear model (Kmenta, Chapter 9, 1971). The bias of the least squares estimator  $\hat{h}(t_n)$  varies with  $n$  unless the covariance of  $\{u(t_n)\}$  is identical to that of  $\{x(t_n)\}$ . The pattern of  $\{\hat{h}(t_n)\}$  will be completely different from the pattern of the true  $h(t_n)$  if the errors are sizable as compared with the true series (Clay and Hinich, 1981).

Assuming some mild restrictions on  $\{h(t_n)\}$  and the additive errors, Clay and Hinich and Hinich (1982) develop consistent estimators of a scalar multiple of the coefficients using a Hilbert transform relationship.



Availability Codes	
Avail and/or	
Special	
A	

This paper considers the discrete-time linear relationship rather than the continuous-time model used in the cited previous work. A Hilbert transform relationship for sequences is used to derive consistent estimators of the  $h(t_n)$  from a consistent estimator of the cross spectrum between  $\{x(t_n)\}$  and  $\{y(t_n)\}$ . The approach uses the frequency domain interpretation of a linear filter (Brillinger, Section 2.7, 1975). The basic terminology necessary to understand the method is reviewed in the next section.

### 1. The Discrete Hilbert Transform

In filter theory terminology,  $\{x(t)\}$  is the input to a linear filter whose impulse response is  $\{h(t_n)\}$ . The filter is called causal if  $h(t_n) = 0$  for  $t_n < 0$ , and thus  $y(t_n)$  depends only on  $x(s_n)$  for  $s_n < t_n$ . In economics, a causal linear model is called a distributed lag with lag parameters  $\{h(t_n)\}$ .

A filter is stable if  $\sum_{n=0}^{\infty} |h(t_n)| < \infty$ . The frequency response of a stable filter is characterized by its transfer function

$$H(f) = \sum_{n=0}^{\infty} h(t_n) \exp(-i2\pi f t_n). \quad (1.1)$$

The inverse relationship is

$$h(t_n) = \tau \int_0^{1/\tau} H(f) \exp(i2\pi f t_n) df. \quad (1.2)$$

Since  $t_n = \tau f$ ,  $H(f + 1/\tau) = H(f)$  for all  $f$ . For real  $\{h(t_n)\}$ ,  $H(-f) = H^*(f)$  where the star denotes complex conjugate, and thus the phase response  $\phi(f) = \tan^{-1}[\text{Im}H(f)/\text{Re}H(f)]$  is an odd function with  $\phi(0) = 0$  for the range  $-\pi < \phi < \pi$ . The filter's gain  $|H(f)|$  is an even function.

The Hilbert transform relates  $\ln|H(f)|$  to  $\phi(f)$  if the causal filter is minimum phase-lag, i.e. if  $H(z) = \sum_{n=0}^{\infty} h(t_n)z^n$  has no zeros on or inside the unit circle  $|z| = 1$  on the complex plane. The phase-lag  $\phi(f)$  of a finite minimum phase-lag filter is less than any other filter with the same gain function (Zadeh and Desoer, Section 9.7, 1963). Normalize the time unit so that  $\tau = 1$ . The log gain is related to its phase by the Hilbert transform

$$\log|H(f)| = \int_0^1 \phi(g) \cot[\pi(f-g)] dg + c, \quad (1.3)$$

where  $c$  is a scaling constant (Gold and Rader, p. 248, 1969). Note that the integrand has a singularity only at  $g = f$  where the principal value of the integral exists.

Suppose that the gain of a minimum phase-lag filter is normalized by setting  $c = 0$ . Then it is proven in the Appendix that  $h(0) = \pm 1$ .

It is clear from (1.3) that accurate estimates of the phases at a dense frequency grid on the unit interval yield accurate estimates of a scalar multiple of the gain at these frequencies using a finite sum approximation to the integral. The estimated gain combined with the estimated phases gives an estimate of the transfer function which can be inverted to yield an estimate of the impulse response (distributed lag) coefficients. If the estimated impulse response is minimum phase, then it follows from the theorem in the Appendix that the estimate of  $h(0)$  is nearly one when  $h(0) = 1$  for the normalization  $c = 0$ , provided the errors are sufficiently small due to the use of a large sample or because the signal-to-noise ratio in the data signals are large.

## 2. Estimating the Phase of the Transfer Function

Accurate estimates of the phases can easily be derived from a consistent estimator of the cross spectrum. Assume that the autocovariance of  $\{x(t_n)\}$  is absolutely summable. Then the spectrum  $S_x(f)$  of  $\{x(t_n)\}$  exists, and more importantly, the cross spectrum is of the form

$$S_{xy}(f) = H(f)S_x(f) \quad (2.1)$$

(Jenkins and Watts, Sect. 8.4.2, 1968). If  $S_x(f) > 0$ , it follows from (2.1) that

$$\phi(f) = \arctan[\text{Im}S_{xy}(f)/\text{Re}S_{xy}(f)]. \quad (2.2)$$

Thus a consistent estimator of  $S_{xy}(f)$  yields a consistent estimator of  $\phi(f)$  using (2.2).

Recall that the sample consists of simultaneous observations of the noisy processes  $\{\tilde{x}(t_n) = x(t_n) + u(t_n)\}$  and  $\{\tilde{y}(t_n) = y(t_n) + \varepsilon(t_n)\}$ . The following assumptions are crucial for the method: 1)  $\{u(t_n)\}$  and  $\{\varepsilon(t_n)\}$  are uncorrelated, and 2) they are uncorrelated with the true processes  $\{x(t_n)\}$  and  $\{y(t_n)\}$ . Given these assumptions, the cross covariance functions of the observed and true processes are identical, i.e.  $c_{\tilde{x}\tilde{y}}(t) = c_{xy}(t)$  for all  $t$ . Thus the cross spectrum of the observed processes is  $S_{\tilde{x}\tilde{y}}(f)$ .

There are several related approaches to estimating the cross spectrum of the observed processes. Consistent estimators and their asymptotic properties are presented by Anderson (Chapter 9, 1971), Brillinger (Chapter 5, 1975), and Fuller (Chapter 7, 1976).

One simple method uses the discrete Fourier transforms of a sample  $\{x(t_n), y(t_n) | n = 0, 1, \dots, N-1\}$ . For frequencies  $f_k = k/N$ , these transforms are:

$$X(f_k) = \sum_{n=0}^{N-1} x(t_n) \exp(-i2\pi nk/N)$$

and (2.3)

$$Y(f_k) = \sum_{n=0}^{N-1} y(t_n) \exp(-i2\pi nk/N).$$

The cross spectrum estimator at frequency  $0 < f_0 < 1/2$  is the average<sup>1</sup>

$$\hat{S}_{xy}^{--}(f_0) = \frac{1}{MN} \sum_{k=-(M-1)/2}^{(M-1)/2} X(f_0 + f_k) Y^*(f_0 + f_k), \quad (2.4)$$

where  $M \ll N$ . From Theorems 7.3.1 and 7.3.2 in Brillinger (1975), the expected value of  $\hat{S}_{xy}^{--}(f_0)$  is  $S_{xy}^{--}(f_0) + O(M/N)$  and its variance is  $O(M^{-1})$ .

Thus, the mean square error is small if and only if  $M$  is large and  $M/N$  is small. To parameterize the tradeoff between bias and variance, let  $M = N^\alpha$  where  $0 < \alpha < 1$ . Good results for the mean square error of the phase of  $\hat{S}_{xy}^{--}(f_0)$  have been obtained in a variety of simulated distributed lag models that have tried using  $M = N^{1/2}$  and  $N = 500$ .

The estimate of phase is given by

$$\hat{\phi}(f_0) = \arctan[\text{Im} \hat{S}_{xy}^{--}(f_0) / \text{Re} \hat{S}_{xy}^{--}(f_0)]. \quad (2.5)$$

Expanding (2.5) in a Taylor series about  $S_{xy}(f_0)$ , it follows that

$\hat{\phi}(f_0) = \phi(f_0) + e$  where the expected value of the error  $e$  is of order  $M/N$ .

For  $1/2 < f_0 < 1$ ,  $\hat{\phi}(f_0) = \hat{\phi}(1-f_0)$ . The variance of the phase error is

$$\sigma_{\hat{\phi}}^2(f_0) = (2M)^{-1} [\gamma^{-2}(f_0) - 1] + O(M^{-2}), \quad (2.6)$$

where

$$\gamma(f_0) = |S_{xy}(f_0)| / [S_x(f_0)S_y(f_0)]^{1/2} \quad (2.7)$$

denotes the coherence between the observed signals.

Since  $M = N^2$ , it follows from (2.6) and the above that  $\hat{\phi}(f_0)$  is a consistent estimator of  $\phi(f_0)$  as  $N \rightarrow \infty$  if  $\gamma(f_0) > 0$ .<sup>2</sup> Also,  $\hat{\phi}(f_j)$  and  $\hat{\phi}(f_k)$  are approximately uncorrelated if  $|j-k| > M$  because the discrete Fourier coefficients have cross correlations of order  $N^{-1}$  if the two time series have well behaved cumulants (Brillinger, Theorem 4.4.2, 1975). If so,  $\hat{\phi}(f_{k(N)})$  is asymptotically Gaussian if  $k(N)/N \rightarrow f_0$  as  $N \rightarrow \infty$ , eg.  $k(N) = [f_0 N]$  where the brackets denote the integer part of the number.

Expression (2.6) can be rewritten in terms of the noise-to-signal ratios for the observed processes. These ratios are denoted  $r_x(f) = S_u(f)/S_x(f)$  and  $r_y(f) = S_e(f)/S_y(f)$ , where  $S_u(f)$  and  $S_e(f)$  are the spectra of the respective additive noises. Then from (2.7),  $\gamma^2(f) = [(1+r_x(f))(1+r_y(f))]^{-1}$ . Thus

$$2M\sigma_{\hat{\phi}}^2(f) = r_x(f) + r_y(f) + r_x(f)r_y(f) + O(M^{-1}), \quad (2.8)$$

which implies that  $\sigma_\phi^2$  is approximately linear in the noise-to-signal ratios when they are small. The coherence becomes small, and  $\sigma_\phi^2$  large, when  $r_x$  and  $r_y$  are big (low signal-to-noise ratios).

### 3. Estimating the Gain

A finite sum approximation to the integral in (1.3) for  $f = k/N$  yields

$$\log|H(f)| = N^{-1} \sum_{\substack{j=0 \\ j \neq k}}^{N-1} \phi(g_j) \cot[\pi(f_k - g_j)] + O(N^{-1}) + c, \quad (3.1)$$

where the sum omits  $j=k$  to avoid the singularity at  $j=k$ . Substituting the estimates  $\hat{\phi}(g_j)$  for phases in (3.1) yields a simple estimate of  $\log|H(f_k)| + c$  which will be denoted  $\text{est}[\log|H(f_k)|]$ . Since the  $\hat{\phi}(g_j)$  are Gaussian and consistent, it follows from (3.1) that  $\text{est}[\log|H(f_{k(N)})|]$  is a consistent Gaussian estimator of  $\log|H(f_0)|$  as  $N \rightarrow \infty$ . The large sample variance of this estimator will now be derived.

Theorem 1. The large sample variance of  $\text{est}[\log|H(f_{k(N)})|]$  is

$$\sigma^2(f_{k(N)}) = \pi^{-2} \sum_{\substack{m=-a+1 \\ m \neq 0}}^b m^{-2} \sigma_\phi^2(f_{k(N)} + g_{mM}), \quad (3.2)$$

where  $a = [f_0 N/M]$  and  $b = [(1-f_0)N/M]$ , and  $g_{mM} = mM/N$ . Thus the large sample standard deviation is of order  $N^{-1/4}$  if  $\alpha=1/2$  (a resolution bandwidth of  $N^{-1/2}$ ).

Proof: The integral in (1.3) can be rewritten as  $-\int_{-f}^{1-f} \phi(f+y) \cot(\pi y) dy$ . A finite sum approximation of this integral that converges as  $M/N = N^\alpha \rightarrow 0$  when  $N \rightarrow \infty$  is

$$-(M/N) \sum_{\substack{m=-a+1 \\ m \neq 0}}^b \phi(f_k + g_{mM}) \cot(\pi g_{mM}) \quad . \quad (3.3)$$

Thus the large sample variance of (3.3) with  $\hat{\phi}$  replacing  $\phi$  is asymptotically equal to the large sample variance of  $\text{est}[\log|H(f_k)|]$ . The correlation between  $\hat{\phi}(mM/N)$  and  $\hat{\phi}(nM/N)$  is  $O(N^{-1})$  for  $m \neq n$ , and thus it follows from (2.6) that the variance of (3.3) for large  $N$  is

$$(M/N)^2 \sum_{\substack{m=-a+1 \\ m \neq 0}}^b \sigma_{\phi}^2(f_k + g_{mM}) \cot^2(\pi g_{mM}) \quad . \quad (3.4)$$

Since  $\cot^2 x = \sin^{-2} x - 1$  and  $N^{-2} \sin^{-2}(\pi mM/N) = (\pi mM)^{-2} + O(N^{-2})$ , for large  $N$  expression (3.4) becomes  $\pi^{-2} \sum_{\substack{m=-a+1 \\ m \neq 0}}^b m^{-2} \sigma_{\phi}^2(f_k + g_{mM})$ , the desired large sample variance  $\sigma^2(f_k)$ .

For example, suppose that the coherence is constant over frequencies. Since  $\sum_{m=1}^{\infty} m^{-2} = \pi^2/6$ , then the large sample variance is  $(\gamma^{-2} - 1)/6M$ . When  $\gamma(f)$  is slowly varying, a useful approximation of  $\sigma^2(f_k)$  is  $[\gamma^{-2}(f_k) - 1]/6M$ .

It will now be shown that the estimates are asymptotically uncorrelated.

Theorem 2. The large sample covariance between  $\text{est}[\log|H(f_{j(N)})|]$  and  $\text{est}[\log|H(f_{k(N)})|]$  for  $j(N) = [f_1 N]$  and  $k(N) = [f_2 N]$  is  $O(1/N|f_2 - f_1|)$ .

Proof: Once again the asymptotic independence of the phase estimators will be used. From (3.3) and (3.4), the large sample covariance is approximated by

$$(M/N)^2 \sum_m \sigma_\phi^2(f_1 + g_{mM}) \cot(\pi g_{mM}) \cot[\pi(g_{mM} - f_2 + f_1)] \quad (3.5)$$

if  $f_1 < f_2$ . The sum excludes the singularities at  $m = 0$ , and  $m = (f_2 - f_1)N/M$  if it is an integer in the set  $\{-[f_1 N/M] < m < [(1-f_2)N/M]\}$ . The largest values of the summand are at  $m = \pm 1$  and  $m = [(f_2 - f_1)N/M] \pm 1$ , where they are approximately  $(M/N)^2 M^{-1} (N/M) / \pi^2 (f_2 - f_1) = [\pi^2 N (f_2 - f_1)]^{-1}$ .

#### 4. Comparison with Other Methods

The numerical approximations of the Hilbert transform described in Clay and Hinich and Hinich can yield very biased estimates in some cases when the sample size is moderate and the bandwidth is large. The method presented here usually gives less finite sample bias than these other methods, especially for discrete-time filters.

As is to be expected however, the large sample variance of  $\text{est}[\log|H(f_k)|]$  is somewhat larger than that of the other estimators. To compare with the older methods, suppose that the cross spectrum is estimated from  $P$  successive samples of size  $N$  (observation time  $N\tau$ ) of  $\{x(t_n), y(t_n)\}$ . The sample cross spectrum at  $f$  is the average  $(p=1, \dots, P)$  of the unsmoothed terms  $\tilde{X}_p(f) \tilde{Y}_p^*(f)$ . If the samples are approximately uncorrelated, the large sample variance of the phase computed from this block smoothed cross spectrum is  $P^{-1} \sigma_\phi^2(f)$  (Hinich and Clay, 1968). Reworking the proof of the theorem using the estimator

$$-N^{-1} \sum_{j=-[f_0 N]+1}^{[(1-f_0)N]} \phi((k+j)/N) \cot(\pi j/N), \text{ its large sample variance is } O(P^{-1}).$$

The large sample variance of the other methods is of order  $o(P^{-1})$  as  $N \rightarrow \infty$ ,

but the computational simplicity and robustness of the discrete-time method makes it much more applicable than the older approaches.

### 5. Estimating the Transfer Function and Impulse Response

Recall that the log gain is estimated up to an additive scale constant. Thus  $\exp\{\text{est}[\log|H(f_k)|]\}$  is a consistent estimator of  $c|H(f_k)|$ . Using the linear approximation of the exponential, the variance of this estimator is approximately  $|H(f_k)|^2 \sigma^2(f_k)$  for large  $M$ . The covariance between this estimator and  $\hat{\phi}(f_k)$  is approximately  $\pi^{-1} \sigma_\phi^2(f_k)$  from (3.3).

The transfer function is estimated by

$$\hat{H}(f_k) = \exp\{\text{est}[\log|H(f_k)|] + i \hat{\phi}(f_k)\}. \quad (5.1)$$

Since the error in the phase estimate is additive, the error in  $\text{est} \log|H(f_k)|$  is additive and thus the error in  $\hat{H}(f_k)$  is multiplicative.

The complex variance of the exponent in (5.1) is defined to be the sum of the variances of the real and imaginary parts, which is  $\sigma^2(f_k) + \sigma_\phi^2 = (2/3M)[\gamma^{-2}(f_k)-1]$  using the simplifying approximation for  $\sigma^2(f_k)$  previously presented. Using the linear approximation of the exponential function, for large  $M$  the complex variance of the estimator is

$$E|\hat{H}(f_k) - H(f_k)|^2 = (2/3M)|H(f_k)|^2[\gamma^{-2}(f_k)-1]. \quad (5.2)$$

From Theorem 2 and the properties of the  $\hat{\phi}(f_k)$ , the complex covariance between  $\hat{H}(f_j)$  and  $\hat{H}(f_k)$  is  $O(1/N|j-k|)$ .

The estimator of the impulse response coefficient  $h(t_n)$  is

$$\hat{h}(t_n) = N^{-1} \sum_{k=0}^{N-1} \hat{H}(f_k) \exp(i2\pi kn/N). \quad (5.3)$$

If the estimated impulse response is minimum phase, then  $\hat{h}(0) = 1 + O(N^{-1})$ .

The error is only due to the errors from approximating the integrals.

For large  $M$ , the variance of  $\hat{h}(t_n)$  is approximated by the variance of  $(M/N) \sum_{m=0}^{[N/M]} \hat{H}(f_{mM}) \exp(i2\pi mM/N)$ , with an error of  $O(M/N)$ . This variance, with an error of  $O(1/NM)$  due to the covariances of the  $\hat{H}(f_{mM})$ , is

$$\begin{aligned} (M/N)^2 \sum_{m=0}^{[N/M]} E |\hat{H}(f_{mM}) - H(f_{mM})|^2 &\approx (2M/3N^2) \sum_{m=0}^{[N/M]} |H(f_{mM})|^2 [\gamma^{-2}(f_{mM}) - 1] \\ &\approx \frac{2}{3N} \int_0^1 |H(f)|^2 [\gamma^{-2}(f) - 1] df. \end{aligned} \quad (5.4)$$

The largest errors in these approximations are  $O(M^{-2})$  or  $O((M/N)^{-2})$ , and thus are  $O(N^{-\beta})$  where  $\beta = 2\min(\alpha, 1-\alpha)$ . Note that this variance is independent of  $n$ , so all parameter estimates have equal variance for large  $N$ .

As an example, suppose that coherence is constant across frequencies. Since  $\int_0^1 |H(f)|^2 df = \sum_{n=0}^{\infty} h^2(t_n)$  by Parseval's formula, assuming existence, then from (2.8) and (5.4) the large sample variance of  $\hat{h}(t_n)$  is

$$\text{Var } \hat{h}(t_n) = \frac{2}{3N} (r_x + r_y + r_x r_y) \sum_{n=0}^{\infty} h^2(t_n). \quad (5.5)$$

Thus  $\hat{h}(t_n)$  is a precise estimator when  $N$  is large and the phase error is small. Expression (5.5) overstates the error in the estimates when the

right hand side is small since it is derived without incorporating the constraint given by  $\hat{h}(0) = N^{-1} \sum_{k=0}^{N-1} \hat{H}(f_k) = \pm 1 + O(N^{-1})$ .

## 6. Data Analysis

The method was tested using artificially created data for the causal minimum phase-lag filter  $\{h(0) = 1, h(1) = 3/2, h(2) = 1, h(3) = 1/2, h(4) = 1/4, h(m) = 0 \text{ for } m > 5\}$ . For each trial a sequence of  $N = 500$  values of  $y(n) = \sum_{m=0}^4 h(m)x(n-m)$  were computed using these five lag parameters and an input sequence generated by the recursion

$$x(n) = 0.5 x(n-1) + z(n), \quad (6.1)$$

where the  $z(n)$  values were generated by the IMSL subroutine GGNML to simulate a sequence of independent normal  $N(0,1)$  variates. In other words, the input to the filter was a pseudo-AR(1) process with parameter  $1/2$  and variance  $4/3$ . Two independent white noise sequences were generated to produce noisy "observations"  $\tilde{x}(n) = x(n) + \sigma u(n)$  and  $\tilde{y}(n) = y(n) + \sigma \epsilon(n)$  where the positive parameter  $\sigma$  controls the standard deviation of the errors. The pseudo-random variates  $u(n)$  and  $\epsilon(n)$  were also  $N(0,1)$  as generated by the GGNML subroutine using different seed values. A variety of seed values were used to test the method for each set of parameter values that were tried.

Table 1 presents typical results for  $\sigma = 0.71, 1.00, 1.22, 1.41,$  and  $1.58$  ( $\sigma^2 = 0.5, 1, 1.5, 2, 2.5$ ) using  $M = 39$  as the smoothing number. The resolution bandwidth is then  $M/N = 0.078$ . The same artificial data was used in each run. Table 1 presents  $\hat{h}(t)$  for  $t = 0, \dots, 6$ ; the root mean square error (RMSE) of  $\hat{h}(t)$  for  $t = 0, \dots, 499$ ; the square root of the

large sample variance (5.4) with the estimated gain and squared coherence in place of their true values (Est SD  $\hat{h}(t)$ ); and the average squared coherence for  $0 < f < 1/2$ . The coherence is largest for low frequencies since  $\{x(t)\}$  is red noise whereas the additive errors are white.

The method produced estimates of the distributed lags that track the impulse response even for the largest value of  $\sigma$  that was used. The estimates of the zero tail of the impulse response are very accurate, especially for the far right tail. For example, the maximum error for  $t > 450$  is 0.0004 for  $\sigma^2 < 1.5$ , 0.003 for  $\sigma^2 = 2$ , and 0.01 for  $\sigma^2 = 2.5$ . This accuracy is greater than that predicted by the asymptotic calculations. When the phase error is small, the length of the estimated impulse response is slightly larger than the true length.

The estimated standard deviation has an upward bias due to biases in the estimated phase and gain. The large sample variance appears to only serve as an upper bound for the sample size  $N = 500$ . The largest errors are in the estimates of  $h(1) = 1.5$ . These estimates are downward biased by the convolution of the true lags with the transform of the errors in the estimated gain. The bias is a function of  $M$ .

Table 2 presents results for  $\sigma = 1.22$  using  $M = 9, 19, 29, 39, 49,$  and 59. The results show how important it is to use different values of  $M$  when using this method to estimate the lags. More research is needed to determine practical ways for setting  $M$  as a function of prior beliefs about the lag pattern, the signal-to-noise ratios, and the shape of the signal spectra.

Table 3 presents ordinary least squares estimates of the  $h(t)$  for  $\sigma^2 = 1.5$  using fifteen lags and an intercept constrained to be zero. The estimates are downward biased as expected. Although the number of non-zero lags is correctly determined by inspection of the  $t$  ratios, the lag pattern is distorted in the OLS estimates.

#### 7. Conclusion

The artificial data results indicate that the Hilbert transform method can become a powerful tool for estimating distributed lags when the signals are observed with additive noise. This estimation procedure can give accurate estimates of the distributed lag coefficients when there is considerable additive noise in the observed time series, provided that the sample size is sufficiently large to yield accurate phase estimates. If the phase of the transfer function can be estimated sufficiently closely, then the impulse response function is determined by the Hilbert transform up to a scalar multiple. The statistical computational aspect of the procedure is simple.

## Footnotes

1. If  $0 < f_0 < M/2N$ , then average the  $X^*(f_0 + f_k)Y(f_0 + f_k)$  for  $0 < f_k < f_0 + M/2N$ . A similar constriction of bandwidth holds when  $1/2 - M/2N < f_0 < 1/2$ .
2. A consistent estimator of  $\gamma(f_0)$  is  $|\hat{S}_{xy}(f_0)| / [\hat{S}_x(f_0)\hat{S}_y(f_0)]^{1/2}$ , where  $\hat{S}_x(f_0)$  and  $\hat{S}_y(f_0)$  are consistent estimators of the two spectra. A prudent policy is to not estimate the phase at  $f_r = r/N$  if its associated estimated coherence is below a preset threshold. The frequencies where the coherence is too low are skipped in the finite sum approximation of the integral in (1.2).

## References

- Anderson, T. W. (1971). The Statistical Analysis of Time Series. Wiley, New York.
- Brillinger, D. (1975). Time Series, Data Analysis and Theory. Holt, Rinehart and Winston, New York.
- Clay, C. S. and M. J. Hinich. (1981). Estimating the Earth's impedance function when there is noise in the electric and magnetic signals. In Second Applied Time Series Symposium, (D. Findley, editor). Academic Press, New York.
- Fuller, W. A. (1976). Introduction to Statistical Time Series. Wiley, New York.
- Gold, B. and C. M. Rader. (1969). Digital Processing of Signals, New York: McGraw-Hill Book Co.
- Hinich, M. J. and C. S. Clay. (1968). The application of the discrete Fourier transform in the estimation of power spectra, coherence, and bispectra of geophysical data. Reviews of Geophysics 6, 347-363.
- Hinich, M. J. (1982). Estimating the gain of a linear filter from noisy data, Handbook of Statistics 3, edited by D. Brillinger and P. R. Krishnaiah.
- Jenkins, G. M. and D. G. Watts. (1968). Spectral Analysis and its Applications. Holden-Day, San Francisco.
- Kmenta, J. (1971). Elements of Econometrics. Macmillan, New York.
- Titchmarsh, E. C. (1952). The Theory of Functions. Oxford, London.
- Zadeh, L. A. and C. A. Desoer. (1963). Linear System Theory. McGraw-Hill, New York.

Table 1

Hilbert Transform Estimates for  $M = 39$ 

$\sigma^2$	0.05	1.00	1.50	2.00	2.50
$\hat{h}(0)$	1.00	1.00	1.00	1.00	0.99
$\hat{h}(1)$	1.33	1.26	1.17	1.14	1.12
$\hat{h}(2)$	0.95	1.01	1.08	1.13	1.17
$\hat{h}(3)$	0.54	0.58	0.57	0.58	0.60
$\hat{h}(4)$	0.28	0.31	0.37	0.41	0.45
$\hat{h}(5)$	0.11	0.16	0.17	0.18	0.19
$\hat{h}(6)$	0.04	0.05	0.05	0.06	0.08
RMSE $\hat{h}(t)$	0.013	0.017	0.021	0.025	0.029
Est SD $\hat{h}(t)$	0.145	0.168	0.185	0.199	0.212
$\bar{\gamma}^2$	0.41	0.31	0.26	0.22	0.20

Table 2  
Results with Different Smoothing Values

$$\sigma^2 = 1.5$$

M	9	19	29	39	49	59
$\hat{h}(0)$	1.00	0.96	1.01	1.00	1.00	1.00
$\hat{h}(1)$	1.12	1.24	1.17	1.17	1.19	1.20
$\hat{h}(2)$	1.14	1.07	1.08	1.08	1.06	1.03
$\hat{h}(3)$	0.66	0.67	0.62	0.57	0.55	0.50
$\hat{h}(4)$	0.38	0.43	0.37	0.37	0.34	0.32
$\hat{h}(5)$	0.11	0.13	0.21	0.17	0.17	0.15
$\hat{h}(6)$	-0.04	0.07	0.01	0.05	0.06	0.10
RMSE $\hat{h}(t)$	0.058	0.046	0.033	0.021	0.023	0.025
Est SD $\hat{h}(t)$	0.192	0.183	0.182	0.185	0.191	0.197
$\bar{\gamma}^2$	0.32	0.28	0.27	0.26	0.26	0.25

Table 3  
 OLS Estimates for Fifteen Lags and a  
 Zero Intercept

Variable	$\sigma^2 = 1.5$		
	Estimate	Standard Error	t Ratio
h(0)	0.53	0.07	7.90
h(1)	0.69	"	9.95
h(2)	0.64		9.25
h(3)	0.46		6.65
h(4)	0.30		4.24
h(5)	0.12		1.78
h(6)	0.04		0.62
h(7)	0.04		0.54
h(8)	0.07		0.94
h(9)	0.03		0.40
h(10)	-0.05		-0.68
h(11)	0.01		0.15
h(12)	-0.05		-0.67
h(13)	-0.06		-0.92
h(14)	-0.03		-0.44
h(15)	0.00		0.04

$R^2 = 0.54$  with and without intercept

## APPENDIX

If  $\int_{t=0}^{\infty} |h(t)| < \infty$ , the analytic continuation of the transfer function  $H(f)$  inside the unit circle  $|z| = 1$  in the complex  $z$ -plane is  $H_c(z) = \int_{t=0}^{\infty} h(t)z^t$ . On the unit circle  $z = \exp(-i2\pi f)$ ,  $H_c(z) = H(f)$ . For  $z = 0$ ,  $H_c(0) = h(0)$ .

If  $H(f)$  is minimum phase,  $H_c(z)$  has no zeros in  $|z| < 1$  and the Hilbert transforms of phase and gain are:

$$\phi(f) = \int_0^1 \log |H(g)| \cot \pi(g-f) dg \quad (A1)$$

and

$$\log |H(f)| = c + \int_0^1 \phi(g) \cot \pi(f-g) dg, \quad (A2)$$

where  $c$  is an arbitrary constant. Set  $c = 0$ .

The following lemmas are used in the proof of the main result.

Lemma 1.  $\int_0^1 \cot \pi(y-x) dy = 0$  for every  $x$ .

Proof: Consider the minimum phase filter defined by  $\{h(0) = 2, h(t) = 0$  for  $t \neq 0\}$ . This filter's transfer function is the constant 2, and thus its phase is zero for all  $f$ . From (A1),

$$0 = \int_0^1 \log |2| \cot \pi(y-x) dy,$$

and the result then follows.

Lemma 2.  $\int_0^1 \log |H(f)| df = 0$  for a minimum phase filter.

Proof: Reversing the order of integration when integrating both sides of (A2),

$$\int_0^1 \log |H(f)| df = \int_0^1 \left[ \int_0^1 \cot \pi (f-g) df \right] \phi(g) dg = 0$$

from Lemma 1.

Theorem.  $h(0) = \pm 1$ .

Proof: Since  $H(z)$  has no zeros or poles in  $|z| < 1$ , by Jensen's theorem (p. 125, Titchmarsh, 1952)

$$\log |H(0)| = \int_0^1 \log |H(f)| df.$$

From Lemma 2,  $\log |H(0)| = 0$  and thus  $\log |h(0)| = 0$ .

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report No. 26	2. GOVT ACCESSION NO. A123460	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Estimating Distributed Lag Coefficients when there are Errors in the Observed Time Series	5. TYPE OF REPORT & PERIOD COVERED Technical Report	
	6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) Melvin J. Hinich	8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0494	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Virginia Polytechnic Institute & State University	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-315	
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Code 436 Statistics and Probability Program Arlington, VA 22217	12. REPORT DATE December 1982	
	13. NUMBER OF PAGES 21	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)  Distribution of this document is unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Distributed Lag, Linear Filter, Hilbert Transform, Phase, Gain, Consistent Estimation		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)		

**FILMED**

**2-83**

**DTIC**