NOAA Technical Memorandum ERL GLERL-85

COVARIANCE PROPERTIES OF ANNUAL NET BASIN SUPPLIES TO THE GREAT LAKES

Steven G. Buchberger

Great Lakes Environmental Research Laboratory
Ann Arbor, Michigan
July 1994

UNITED STATES
DEPARTMENT OF COMMERCE

Ronald H. Brown
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION

D. James Baker
Under Secretary for Oceans
and Atmosphere/Administrator

Environmental Research
Laboratories

Alan R. Thomas
Director

CONTENTS

TABLES

NOTATION

A_i	dimensionless variance ratio for lake i defined in equation (14a)
B_i	dimensionless variance ratio for lake i defined in equation (14b)
C_i	dimensionless term for lake i defined in equation (20)
$D_i(t)$	annual interbasin diversion at lake i in year t
$E_i(t)$	annual evaporation from lake i in year t
$F(i,j)$	sum of absolute differences between sample and predicted CCFs
$G_i(t)$	annual groundwater flow out of lake i in year t
$N_i(t)$	annual net basin supply to lake i in year t
$P_i(t)$	annual precipitation onto lake i in year t
$Q_i(t)$	annual connecting channel outflow from lake i in year t
$R_i(t)$	annual basin runoff into lake i in year t
$S_i(t)$	annual rate of storage in lake i in year t
$a_i(t)$	noise term in AR(1) model of annual lake outflows
$b_i(t)$	noise term in AR(1) model of annual lake storage rates
$e_i(t)$	residual term in linear reservior model
k	annual lag (years)
$r_{N_i,N_j}(k)$	sample CCF between annual net basin supplies to lakes i and j
$\gamma_{N_i,N_j}(k)$	covariance function for net basin supplies to lake i and lake j.
$\rho_{N_i,N_l}(k)$	predicted CCF between annual net basin supplies to lakes i and j
α_i	AR(1) parameter for annual outflow from lake i
β_i	AR(1) parameter for annual storage in lake i
λ_i	linear reservoir storage coefficient for lake i
ω_i	weighting coefficient for lake i
μ	mean
σ^2	variance
σ_R^2	dimensionless variance ratio, $\sigma_{\Delta Q}^2/\sigma_{\Delta S}^2$, defined in equation (15)
$\pi_{i,j}$	dimensionless term defined in equation (7)

COVARIANCE PROPERTIES OF ANNUAL NET BASIN SUPPLIES TO THE GREAT LAKES

Steven Buchberger[1]

ABSTRACT. The cross correlation function (CCF) and autocorrelation function (ACF) for Great Lakes annual net basin supplies are derived under the assumptions that annual lake outflows and water levels are autoregressive lag-l processes, and that each lake behaves as a linear reservoir. Except for the pairing between Lakes Superior and Michigan-Huron, there is reasonable agreement between the sample and predicted CCFs, especially among net basin supplies to the small lower basin lakes. The derived ACFs reduce to an expression identical in form to the ACF for an Auto Regressive Moving Average (ARMA)(1 ,1) process at all the Great Lakes except Lake Michigan-Huron. At the upper basin lakes, sample and predicted ACFs drop to zero rapidly. At the lower basin lakes, the ACFs exhibit a much more gradual decay suggesting the presence of long-term persistence. Prominent tails in the ACFs of the annual net basin supplies have been attributed to historical shifts in the precipitation regime at the lower basin lakes. Results from this study show that the residual method currently used to estimate net basin supplies can also induce a similar artificial long tail in the ACE This observation has important ramifications in efforts to simulate Great Lakes water levels, since simulation results are quite sensitive to the covariance structure of the annual net basin supplies.

1. INTRODUCTION

Water moving through the Great Lakes is classified according to two primary pathways: (1) inter-basin flows, or (2) net basin supplies. Inter-basin flows represent water conveyed through the natural connecting channels and water imported or exported through man-made diversions. Net basin supply is a derived quantity used to account for all other processes by which water enters or leaves a lake. Included here are water gains due to over-lake precipitation and basin runoff and water losses resulting from lake evaporation and groundwater outflow.

Net basin supplies are useful in short-term water level forecasting, long-term simulation of lake levels and other investigations of Great Lakes hydrology (Croley and Hartmann, 1987). Several multivariate autoregressive moving average (ARMA) time series models have been developed to simulate and forecast monthly net basin supplies (Yevjevich, 1975; Loucks, 1989; Buchberger, 1992; Rassam et al., 1992). These ARMA models can be classified as either direct or indirect. The direct approach considers only monthly net basin supplies and, as a consequence, fails to capture the year-to-year correlations among annual net basin supplies. In contrast, the indirect approach first generates annual net basin supplies and then disaggregates these values into monthly supplies. The disaggregation scheme is able to reproduce both the annual and monthly correlations among the net basin supplies.

GLERL Contribution No. 899
[1] Dept. of Civil and Environmental Engineering, Univ. of Cincinnati, Cincinnati, OH 45221-007 1

Engineered outlet works have regulated the release from Lake Superior since 1921 and from Lake Ontario since 1958. Although discharges from the other three Great Lakes are not directly regulated by man-made control structures, outflows from the middle lakes have been influenced over time by dredging and other modifications to the connecting channels (Quinn, 1985).

2.2 Net Basin Supplies

Let $N_i(t)$ denote annual net basin supply to lake i during year t. By definition, $N_i(t)$ is the sum of annual water gains due to overlake precipitation $P_i(t)$ and tributary runoff $R_i(t)$ minus annual water losses to lake evaporation $E_i(t)$ and seepage $G_i(t)$ or

$$N_i(t) = P_i(t) + R_i(t) - E_i(t) - G_i(t). \tag{1}$$

Owing to the large surface and drainage areas of the Great Lakes, the individual components in equation (1) are difficult to measure. In practice, $N_i(t)$ is usually estimated as the residual term in the lake water balance equation,

$$N_i(t) = Q_i(t) - Q_{i-1}(t) + S_i(t) - S_i(t-1) \pm D_i(t) \tag{2}$$

where Q,,(t) and Q,(t) are, respectively, annual inflows and outflows through the connecting channels at lake i during year t, $S_i(t)$-$S_i(t-1)$ is the corresponding annual change in lake water storage during year t expressed as a rate, and $D_i(t)$ is the annual diversion accounting for water imports and exports at the lake.

The residual approach for estimating $N_i(t)$ is expedient, but it has some shortcomings. The magnitudes of the individual Q_i and S_i terms on the right hand side of (2) are often significantly greater than the resulting N_i on the left hand side. Hence, the residual method is susceptible to significant computational errors introduced by deriving the net basin supply as the difference between large uncertain quantities. For example, Quinn and Guerra (1986) have shown that a 5% error in the estimated flows of the Detroit or Niagara Rivers leads to a 34% error in the computed net basin supply for Lake Erie. Croley (1987) shows that changes in lake storage are especially difficult to quantify during the fall and winter months when frequent storm activity perturbs the lake-wide average water surface. In addition, the residual approach ignores thermal volumetric changes, an omission which can lead to significant errors in the estimated net basin supply (Croley and Lee, 1993). Loucks (1989) provides a good discussion on the advantages and disadvantages of the two basic methods, given in equations (1) and (2), for computing net basin supplies.

2.3 Historical Data

Water levels and outflows of the Great Lakes represent one of the longest and most complete hydrologic records in North America. For this study, annual outflow rates in cubic meters per second and January 1 lake elevations in meters for the period 1900-1989 were obtained from the Great Lakes Environmental Research Laboratory (Hunter and Croley, 1991). Values of coordinated monthly net basin supplies computed with the residual method (equation 2) for the period 1900-1989 were provided by the U.S. Army Corps of Engineers. These coordinated monthly net basin supplies are identical to those used in the recent levels reference study prepared for the International Joint Commission (Working Committee 3, 1993). The term "coordinated" implies that these net basin supply data have been published under the auspices of the bi-national coordinating committee on Great Lakes basic hydraulic and hydrologic data.

Table 2a.--Properties of Great Lakes Annual Storage Rates, 1900- 1989.

Statistic	Superior	Michuron	St. Clair	Erie	Ontario
Mean (m^3/s)	7,891	22,929	165.6	3,068	2,647
Standard Dev (m^3/s)	368.3	1,336.2	13.64	266.7	163.6
Skewness	-0.369	-0.233	0.178	0.334	-0.269
ACF lag-1 Coefficient	0.512	0.797	0.675	0.807	0.579
Coefficient of Variation	0.047	0.058	0.082	0.087	0.062

Table 2b.--Lag-Zero Cross-Covariances and Cross-Correlations Among Great Lakes Annual Storage Rates, 1900-1989.

Lake	Superior	Michuron	St. Clair	Erie	Ontario
(1) Superior	135,640 (1.000)	237.124 (0.482)	3,109 (0.605)	41,781 (0.425)	16,119 (0.268)
(2) Michigan-Huron		1,785,440 (1 .000)	16,319 (0.856)	310,958 (0.873)	147.685 (0.676)
(3) St. Clair			186.2 (1.000)	3,436 (0.900)	1,537 (0.674)
(4) Erie				71,106 (1.000)	31,835 (0.730)
(5) Ontario					26,758 (1.000)

Notes: (1) Table 2b is symmetric.
(2) Year starts on January 1.
(3) Crosscovariance units are (cubic meters per second)^2.
(4) Cross-correlations are given in parenthesis.
(5) St. Clair data from period 1910 to 1989.

reasonable for the Great Lakes where surface areas are large and lake level fluctuations are small. The surface area, conversion factor and elevation datum for each lake are summarized in Table 4. In what follows, annual water levels and annual storage rates are used interchangeably since they are related to each other by a lake specific conversion factor.

The mean storage rates for each lake given in Table 2a depend on the datum selected. This has no bearing on the covariances listed in Table 2b, since the covariances of the storage rates are independent of the datum selected. Because storage rates depend on lake size, the covariances of the annual storage rates given in Table 2b take their maximum values at Michigan-Huron (the largest lake) and their minimum values at St. Clair (the smallest lake). To get the statistics in Tables 3a and 3b, coordinated monthly net basin supplies were converted to an equivalent annual net basin supply also expressed as cubic meters per second.

The outflows and storage rates at the Great Lakes have high temporal and spatial correlations. As shown in Tables la and 2a, the lag-l autocorrelations in the annual outflows and the storage rates range from 0.5 1 to 0.85. Similarly, the lag-0 cross correlations among annual outflows given in Table lb range from about 0.33 (Superior and the lower lakes) to values exceeding 0.98 (Michigan-Huron and St. Clair; Erie and Ontario). The lag-0 cross correlations among annual storage rates shown in Table 2b range from about 0.27 (Superior and Ontario) to 0.90 (St. Clair and Erie). Without exception, the lowest observed values of the temporal and spatial correlations occur at the top of the chain with Lake Superior (see row 1 in Tables lb and 2b). In part, this may reflect the effects of regulation plans that have controlled outflows and water levels at Lake Superior since 1921. Although not quite as dramatic as the lake outflows and storage rates, annual net basin supplies have lag-l autocorrelations given in Table 3a that range from 0.16 (Superior) to 0.50 (St. Clair) and lag-0 cross correlations given in Table 3b ranging from 0.25 (Superior and St. Clair) to 0.66 (Erie and Ontario).

3.0 MODELING STRATEGY

3.1 ARMA Models

How does the presence of significant temporal and spatial correlations among the lake outflows and water levels affect the covariance structure of annual net basin supplies that are computed with the residual method? To examine this issue, autocorrelation and cross correlation structures for the annual outflows and lake levels must be specified. Exploratory data analyses showed that three **ARMA** type models-AR(l), **AR(2),** and ARMA(**1,1**)-provide adequate descriptions of the autocorrelation in these annual series. As summarized in Table 5, no single model consistently emerged as the top candidate. For annual lake outflows, the ARMA(**1,1)** model tended to give the best fit on the upper lakes while the AR(1) model prevailed on the lower lakes. For annual lake levels, the AR(1) model tended to work best, except on Lake St. Clair.

In the presence of several competing candidates, here we adopt the AR(1) option as the simplest model which offers the most universal application to annual outflows and water levels throughout the entire Great Lakes system. There is some precedent here. An AR(1) model has been suggested by Potter (1992) for annual water levels in Lake Erie and by Yevjevich (1972) for annual outflows from Lake Ontario. Letting Q*(r) = $Q_i(t) - \mu_{Q_i}$ and $S_i^*(t) = S_i(t) - \mu_{S_i}$ where μ_{Q_i} is the mean annual outflow from lake i, and μ_{S_i} is mean annual storage rate in lake **i,** then the **AR(1)** models for lake outflows and storage rates are written

$$Q_i^*(t) = \alpha_i Q_i^*(t-1) + a_i(t) \tag{3a}$$

$$S_i^*(t) = \beta_i S_i^*(t-1) + b_i(t) \tag{3b}$$

Table X--Ranking of **ARMA** Models Fitted to Great Lakes Annual Outflows and Water Levels, **1900-1989.**

Data Series	Superior	Michuron	St. Clair	Erie	Ontario
Annual Lake outflows	**ARMA(1,1)** **AR(1)** **AR(2)**	**ARMA(1,1)** **AR(2)** **AR(I)**	**ARMA(1,1)** **AR(2)** **AR(1)**	**AR(1)** **ARMA(1,1)** **AR(2)**	**AR(1)** **ARMA(1,1)** **AR(2)**
Annual **Lake** Levels	**AR(1)** **AR(2)** **ARMA(1,1)**	**AR(1)** **AR(2)** **ARMA(1,1)**	**AR(2)** **ARMA(1,1)** **AR(l)**	**AR(1)** **AR(2)** **ARMA(1,1)**	**AR(1)** **AR(2)** **ARMA(1,1)**

Notes: (1) Model ranking from top (best) to bottom (worst) is based on **Akaike** Information Criteria **(Akaike,** 1974).

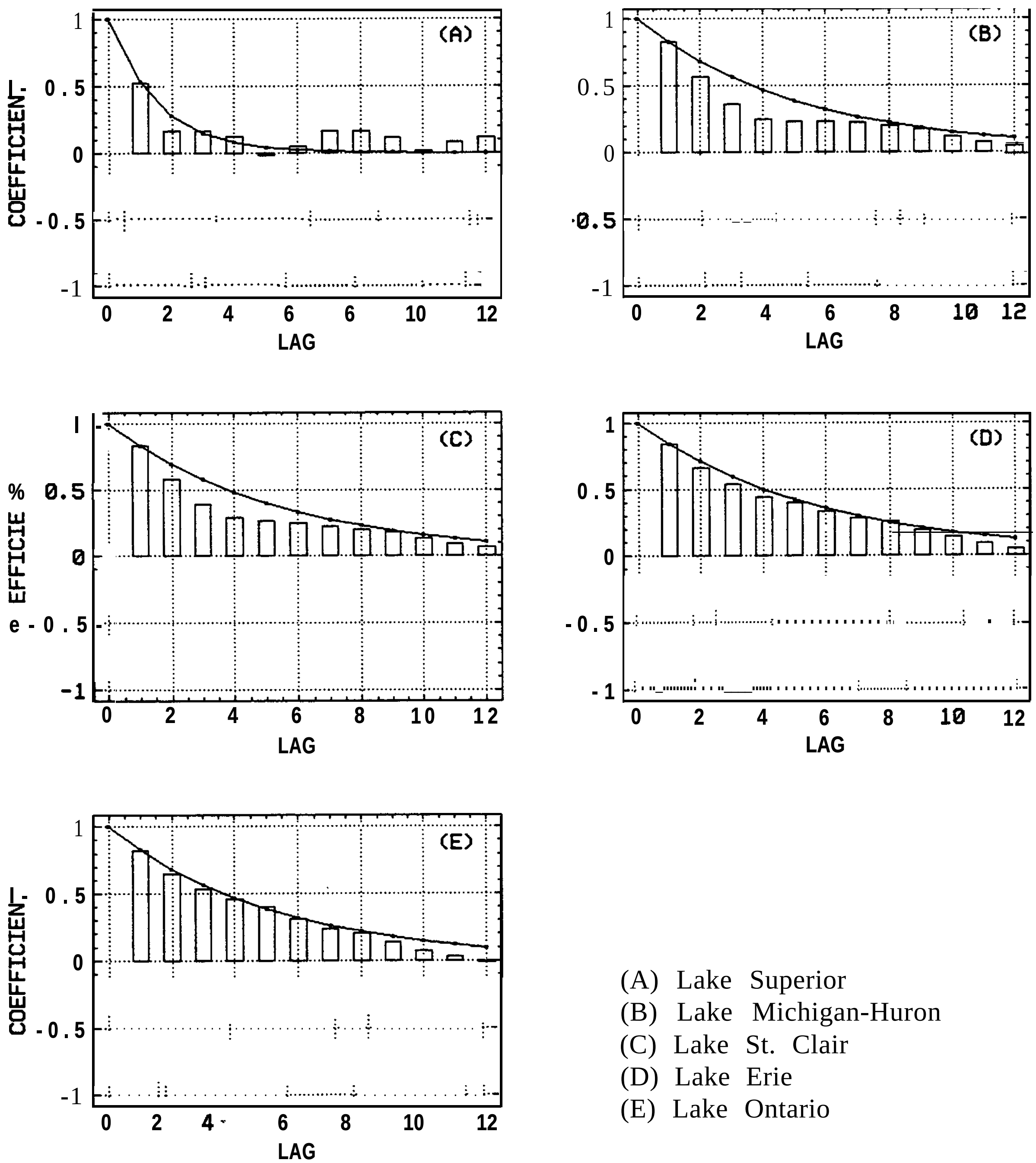


Figure 2.--Sample and AR(1) ACFs for Great Lakes annual connecting channel flows.

An ideal linear reservoir has storage proportional to outflow, $S(t)=\lambda Q(t)$ where λ is the storage coefficient (Klemes, 1974a). Here we adopt a discretized form of the linear reservoir model to better reflect the Great Lakes data used in this study

$$S_i(t) = \lambda_i\left[\omega Q_i(t) + (1-\omega)Q_i(t+1)\right] + e_i(t) \tag{4}$$

In equation (4) $0 \le \omega \le 1$ is a factor used to express lake storage as a weighted average of annual lake outflows, and $e_i(t)$ is an error term with zero mean representing the difference between observed storage levels and the predicted linear reservoir response. If $\omega=0$, then equation (4) gives a "controlled" linear reservoir in the sense that the immediately past storage level governs the current output (Klemes, 1974a). If $\omega=1$, equation (4) gives a "spontaneous" linear reservoir where the current outflow depends on the current storage level and vice versa. Here we take $\omega=0.50$ to give equal emphasis to the annual outflows that occur during the period immediately preceding and following the January 1 lake level, as shown in Figure 4. With ω specified, the storage coefficient λ is found from a linear regression of S(t) on $1/2[Q(t)+Q(t+1)]$. Estimates of λ and the corresponding correlation coefficient $r_{Q,S}$ are given in Table 6. Plots of the regression lines, given in Figure 5, show that the linear reservoir model works particularly well at the middle lakes where outflows are not regulated.

3.3 Covariance Functions

From equation (2), the lag-k covariance between net basin supplies for lakes i and j can be written

$$Cov\left[N_i(t), N_j(t+k)\right] = Cov\left[\left\{\Delta Q_i(t) + \Delta S_i(t)\right\}, \left\{\Delta Q_j(t+k) + \Delta S_j(t+k)\right\}\right] \tag{5}$$

where $\Delta Q_i(t) = Q_i(t) - Q_{i-1}(t)$ and $\Delta S_i(t) = S_i(t) - S_i(t-1)$. Letting $\gamma_{N_i,N_j}(k) = Cov\left[N_i(t), N_j(t+k)\right]$, then as shown in the Appendix, equation (5) leads to

Case 1: k = 1,2,...

$$\begin{aligned}\gamma_{N_i,N_j}(k) &= \alpha_j^k(1+\pi_{j,i}-\pi_{j,j})\gamma_{Q_i,Q_j} - \alpha_j^k(1-\pi_{j,j})\gamma_{Q_{i-1},Q_j} \\ &-\alpha_{j-1}^k(1+\pi_{j-1,i})\gamma_{Q_i,Q_{j-1}} + \alpha_{j-1}^k\gamma_{Q_{i-1},Q_{j-1}} - \beta_j^{k-1}(1-\beta_j)^2\gamma_{S_i,S_j}\end{aligned} \tag{6a}$$

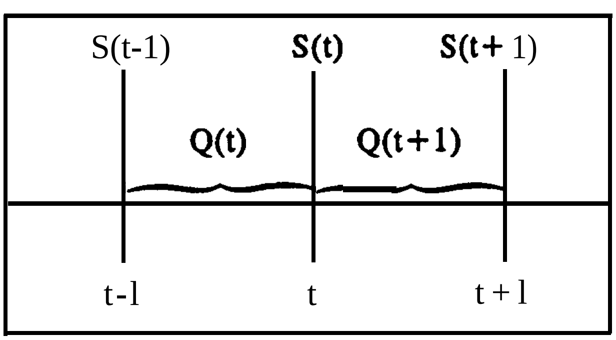


Figure 4.--Discrete time representation of Great Lakes annual connecting channel flows and January 1 storage levels.

Case2: k = 0

$$\gamma_{N_i,N_j}(0) = \left[1 + \frac{(\lambda_i - \lambda_j)}{2}(\alpha_i - \alpha_j)\right]\gamma_{Q_i,Q_j} - \left[1 + \frac{\lambda_i}{2}(\alpha_i - \alpha_{j-1})\right]\gamma_{Q_i,Q_j}$$
$$-\left[1 + \frac{\lambda_j}{2}(\alpha_i - \alpha_{i-1})\right]\gamma_{Q_{i-1},Q_j} + \gamma_{Q_{i-1},Q_{j-1}} + (2 - \beta_i - \beta_j)\gamma_{S_i,S_j} \tag{6b}$$

Case 3: **k = -1,-2,...**

$$\gamma_{N_i,N_j}(k) = \alpha_i^{|k|}(1 + \pi_{i,j} - \pi_{i,i})\gamma_{Q_i,Q_j} - \alpha_i^{|k|}(1 - \pi_{i,i})\gamma_{Q_i,Q_{j-1}}$$
$$-\alpha_{i-1}^{|k|}(1 + \pi_{i-1,j})\gamma_{Q_{i-1},Q_j} + \alpha_{i-1}^{|k|}\gamma_{Q_{i-1},Q_{j-1}} - \beta_i^{|k|-1}(1 - \beta_i)^2\gamma_{S_i,S_j}. \tag{6c}$$

To simplify notation, lag-0 cross-covariance terms have been abbreviated so that $\gamma_{X,Y}$ implies $\gamma_{X,Y}(0)$. The $\pi_{i,j}$ terms appearing in equations (6a) and (6c) are defmed by

$$\pi_{i,j} = \left(\frac{1 - \alpha_i^2}{2\alpha_i}\right)\lambda_j \tag{7}$$

Comparing results for the positive and negative lags shows

$$\gamma_{N_i,N_j}(-k) = \gamma_{N_j,N_i}(k)$$

which is consistent with known cross covariance properties (Box and Jenkins, 1976).

3.4 Cross Correlation Function

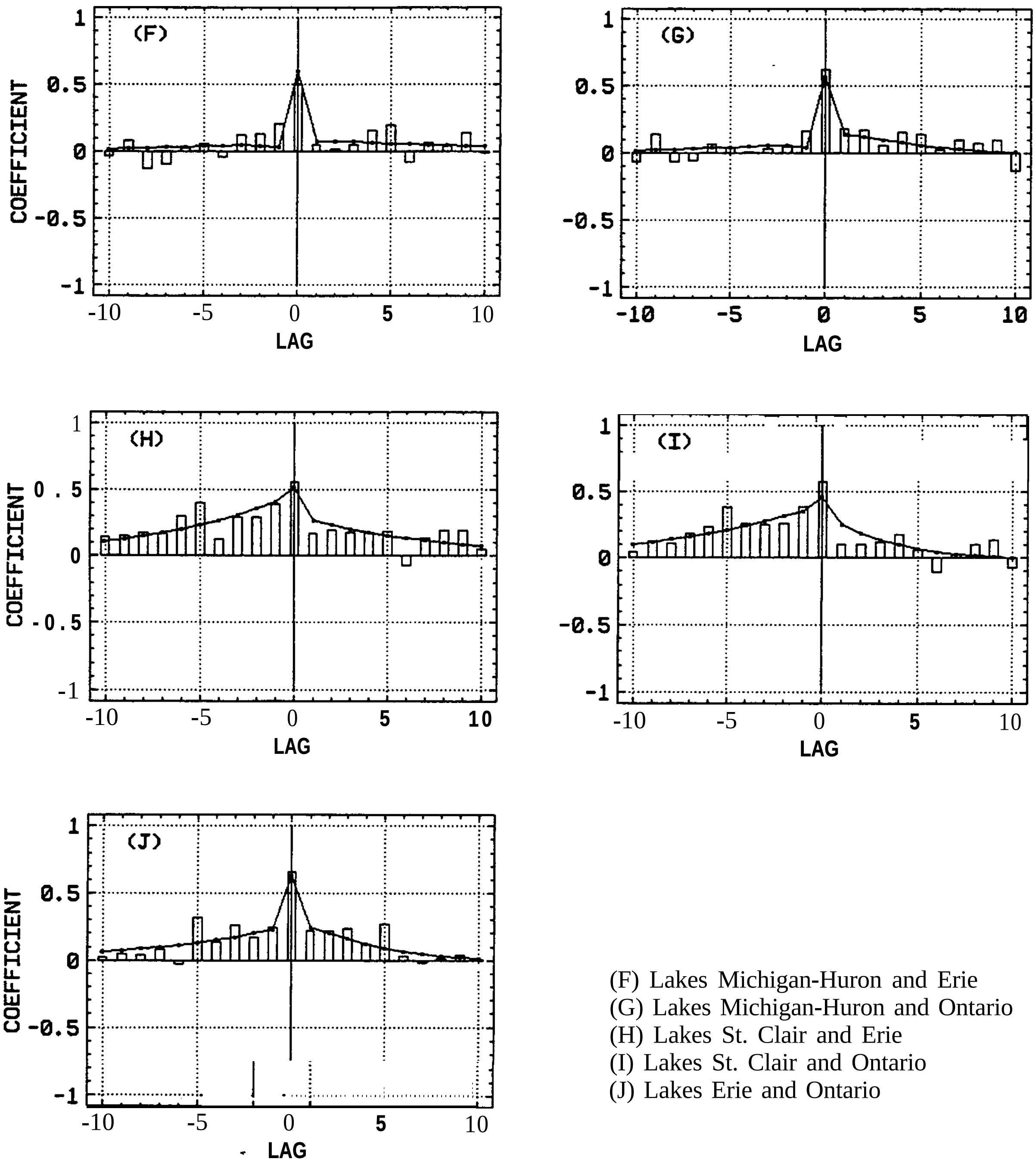
The cross correlation function (CCF) between annual net basin supplies for lakes i and j is defined by

$$\rho_{N_i,N_j}(k) = \frac{\gamma_{N_i,N_j}(k)}{\sigma_{N_i}\sigma_{N_j}} \qquad k = 0, \pm 1, \pm 2, \ldots \tag{8}$$

where σ_{N_i} is the standard deviation of the annual net basin supply for lake i. From the result in (6b), the lag-0 covariance $\gamma_{N_i,N_i}(0) = \sigma_{N_i}{}^2$ is given by

$$\sigma_{N_i}^2 = \sigma_{Q_i}^2 - \left[2 + (\alpha_i - \alpha_{i-1})\lambda_i\right]\gamma_{Q_i,Q_{i-1}} + \sigma_{Q_{i-1}}^2 + 2(1 - \beta_i)\sigma_{S_i}^2. \tag{9}$$

Using expressions (6) and (9) in equation (8) gives the CCF for the annual net basin supplies of any two lakes. Observed and predicted CCFs for annual lags ranging from -10 to +10 at each pair of lakes are presented in Figure 6.



(F) Lakes Michigan-Huron and Erie
(G) Lakes Michigan-Huron and Ontario
(H) Lakes St. Clair and Erie
(I) Lakes St. Clair and Ontario
(J) Lakes Erie and Ontario

Figure 6 (cant).--Sample and predicted CCFs for Great Lakes annual net basin supplies.

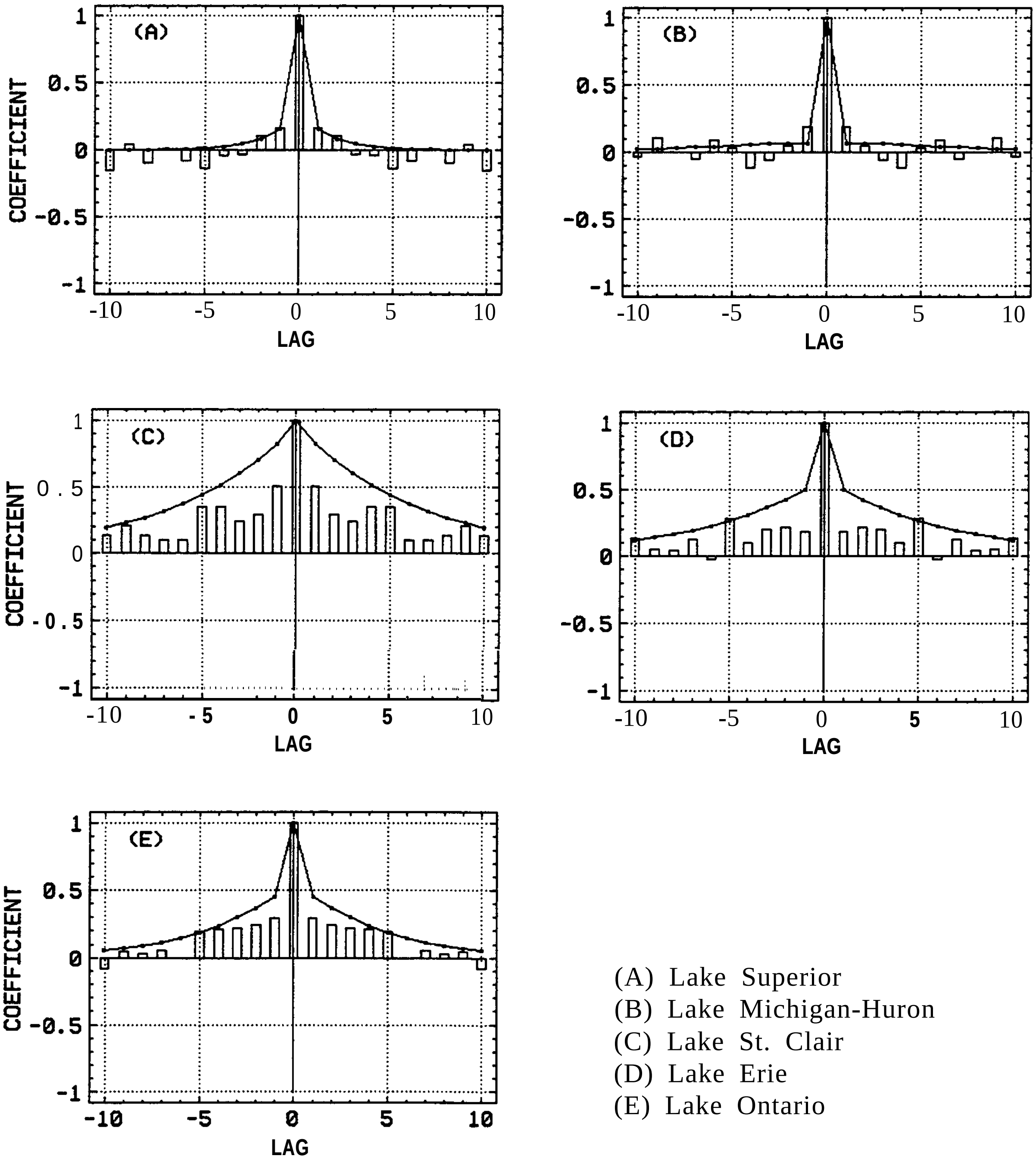


Figure 7.--Sample and predicted ACFs for Great Lakes annual net basin supplies.

Table 7.--Sum of Absolute Differences Between Observed and Predicted Cross Correlations for Great Lakes Annual Net Basin Supplies, 1900- 1989.

Lake	Superior	Michuron	St. Clair	Erie	Ontario
(1) Superior	1.470 (*,6)	2.417 (10,13)	1.090 (2,2)	1.653 (9,12)	1.629 (8,10)
(2) Michigan-Huron		1.509 (*,7)	1.604 (7,9)	1.573 (6,8)	1.342 (4,4)
(3) St. Clair			4.102 (*,15)	1.292 (3,3)	1.357 (5,5)
(4) Erie				2.910 (*,14)	1.071 (1,1)
(5) Ontario					1.642 (*,11)
Rank Sums					
Excluding ACF:	29	27	17	19	18
Including ACF:	43	41	34	38	31

Notes: (1) Table 7 is symmetric.
(2) First entry in parenthesis is rank excluding diagonal element (ACF).
(3) Second entry in parenthesis is rank including diagonal element (ACF).

4.2 Autocorrelation Function

Observed and predicted ACFs for the annual net basin supplies are shown in Figure 7. In contrast to the CCFs, the ACFs show better agreement at the large upper basin lakes than at small lower basin lakes. What is most interesting though is the behavior of both the sample and predicted ACFs. It is clear that the correlation structure of the annual net basin supplies differs between the upper basin lakes and the lower basin lakes.

At Lakes Superior and Michigan-Huron, the ACF quickly dies out with increasing lags. No ACF values for lags beyond $k=0$ are significant at the 95% level. On Lakes St. Clair, Erie and Ontario, both the sample and the predicted ACFs show a much more gradual decay. At St. Clair, the sample ACF is significant at 4 of the first 5 lags beyond k=O. The ACFs for Lakes Erie and Ontario each have significant sample estimates at 2 of the first 5 lags. The match between sample and predicted ACFs is not particularly good at Lakes St. Clair and Erie. This is born out in Table 7 where the absolute difference criterion given by equation (16) takes its maximum values at $F(3,3)=4.102$ and $F(4,4)=2.910$. Nonetheless, the predicted ACFs correctly identify the lakes where the tails of the sample ACFs are most prominent.

As shown in equation (13), the ACF of the annual net basin supplies at all of the Great Lakes except Michigan-Huron can be estimated as the difference between two terms that follow a geometric decay for $|k|\geq 1$. When applied on an individual lake basis, using parameter values from Table 6, equation (13) leads to the following approximations for the ACF of annual net basin supplies for $k=\pm 1, \pm 2, \ldots$:

$$\textit{Superior}, \alpha_1 \approx \beta_1: \quad \rho_{N_1}(k) \approx (A_1 - B_1)(\alpha_1)^{|k|} = (0.28)(0.52)^{|k|} \tag{17a}$$

$$\textit{Michigan - Huron}: \quad \rho_{N_2}(k) \textit{ from equation } (10) \text{ only} \tag{17b}$$

$$\textit{St.Clair}, B_3 \approx 0: \quad \rho_{N_3}(k) \approx (A_3)(\alpha_3)^{|k|} = (0.97)(0.83)^{|k|} \tag{17c}$$

$$\textit{Erie}, \alpha_4 \approx \beta_4, B_4 \approx 0: \rho_{N_4}(k) \approx (A_4 - B_4)(\alpha_4)^{|k|} = (0.59)(0.82)^{|k|} \tag{17d}$$

$$\textit{Ontario}, \beta_5 \approx 0: \quad \rho_{N_5}(k) \approx (A_5)(\alpha_5)^{|k|} = (0.65)(0.83)^{|k|} \tag{17e}$$

meters per year at Lake Superior up to 0.31 meters per year at Lake St. Clair. The relative variation in $\sigma_{\Delta h}$ is small, however, compared to the range in AL which starts at 1,110 km^2 with Lake St. Clair and exceeds 117,000 km^2 at Lake Michigan-Huron. Hence, lake size is the dominant factor in determining the magnitude of $\sigma_{\Delta S}$ in the Great Lakes. For the net connecting channel flows, note that the numerator in equation (15) can be written

$$\sigma^2_{\Delta Q_i} = \sigma^2_{Q_i} + \sigma^2_{Q_{i-1}} - 2\rho_{i,i-1}\sigma_{Q_i}\sigma_{Q_{i-1}} \tag{21}$$

where $\rho_{i\,i-1}$ is the correlation between Q_i and Q_{i-1}. The entries in Table lb show that $\rho_{i,i-1}$ decreases as lake size increases. For example, at Lake St. Clair, $\rho_{3,2}$=0.995 which indicates the annual inflows and outflows are nearly perfectly correlated at the smallest lake in the chain. By virtue of equation (21), high positive correlation between inflows and outflows strongly limits the variability in their difference. In contrast, at Lake Michigan-Huron, $\rho_{2,1}$=0.403 revealing a weaker association between annual inflows and outflows. Here the covariance term has little influence in limiting the variability of the difference between the inflows and outflows at the largest lake. From a physical point of view, this behavior seems reasonable. Small lakes have short detention times and little capacity to moderate input; hence, inflow soon becomes outflow, and the correlation between the two will tend to be high. The converse holds for large lakes.

The observation that both $\sigma_{\Delta Q}$ and $\sigma_{\Delta S}$ increase with lake size in the Great Lakes system probably applies to other lakes in series, though it would not be difficult to devise exceptions to this behavior. From the standpoint of the implicit ARMA(1,1) moving average parameter given by (19), it is the variance ratio $\sigma_R^2=\sigma_{\Delta Q}^2/\sigma_{\Delta S}^2$ which is important in the Great Lakes. Since $\sigma_{\Delta S}^2$ grows faster than does $\sigma_{\Delta Q}^2$, the variance ratio σ_R^2 actually *decreases* with increasing lake size as shown in Table 6. Evidently, a unit increase in lake surface area contributes incrementally more to the variability in lake storage than to the variability in the net connecting channel flow at the Great Lakes. These findings can now be used to help interpret the ACFs given in Figure 7.

In the upper basin, both the sample and predicted ACFs of the annual net basin supplies decay quickly. At Lake Superior, the magnitude of the autoregressive parameter (α_1=0.530) is too small to sustain the tail in the ACF needed for long-term persistence. This is evident in Figure 7-A where both the sample and the predicted ACFs die out after 2 lags. At Lake Michigan-Huron, the autoregressive parameter (α_2=0.827) is suffkiently high to generate long-term persistence with an ARMA(1,1) model. However, the derived ACF does not reduce to au ARMA(1,1) form since $\alpha_2 \neq \alpha_1$. Here too, the sample ACF dies out after only 1 lag as shown in Figure 7-B. Maximum likelihood estimates of the ARMA(1,1) parameters ϕ_{ML} and ϕ_{ML} for the upper basin lakes, listed in Table 6, are not significantly different than zero.

At Lake St. Clair, the variance ratio ($\sigma_{R_3}{}^2$ = 30.0) takes its maximum value. Using equations (19) and (20b), this leads to a very low value for the implicit moving average parameter (θ_3=0.076). In conjunction with the high autoregressive parameter (α_3=0.832), this produces an ACF with very prominent tails as seen in Figure 7-C. Even though the predicted ACF overestimates the sample ACF, it nonetheless correctly identifies St. Clair as the lake where the ACF signature of the annual net basin supplies is most pronounced.

Lakes Erie and Ontario in the lower basin also exhibit prominent positive tails in their sample and predicted ACFs as shown in Figures 7-D and 7-E. Although the predicted ACFs tend to overestimate the observed ACFs, both exhibit a similar pattern, especially at Lake Ontario. Values of the autoregressive parameters (α_4=0.841; α_5=0.826), variance ratios ($\sigma_{R_4}{}^2$ = 1.706; $\sigma_{R_5}{}^2$ = 1.836), and resulting implicit here that the residual method for estimating net basin supplies leads to an autocotrelation structure in the resulting annual. time series that is identical to an ARMA(1,1) model. Provided that the autoregressive parameter is not too small, the derived ACFs exhibit long tails. The strength of this apparent persistence is governed by a dimensionless variance ratio σ_R^2 which increases as lake size decreases. This suggests that the covariance properties of annual net basin supplies estimated from the residual method have an implicit dependence on lake scale.

Certainly, models are available (Salas et al., 1985) that can preserve the annual covariance structure observed in the historical net basin supplies. A key issue in selecting and calibrating these models is to identify the appropriate degree of annual correlation that should be maintained. The challenge introduced here is to reconcile apparent persistence that arises as an artifact of the computational method with genuine persistence that accompanies the behavior of nonstationary physical processes. This has important implications for future modeling efforts since recent studies have shown that simulated Great Lakes water levels are quite sensitive to the presence of autocotrelation in the annual net basin supplies, especially at the lower basin lakes (Rassam et al., 1992). It is reasonable to speculate that the observed high lag autocorrelations and cross correlations in the historical annual net basin supplies at the lower lakes likely reflect the combined effects of artificial and genuine persistence. The only sure way to isolate these potential sources is to estimate net basin supplies directly using equation (1) with accurate observations of water fluxes at many stations providing good spatial coverage for long periods of time.

6.0 CONCLUSIONS

This study examines how temporal and spatial correlations among the connecting channel flows and the lake water levels affect the covariance structure of the Great Lakes annual net basin supplies when these supplies are computed as the residual term in a lake water balance. Results demonstrate that the residual method introduces an autocovariance structure identical in form to the autocovariance of an ARMA(1,1) model at all lakes except Michigan-Huron. The autoregressive parameter ϕ of the ARMA(1,1) model for net basin supplies is equal to the autoregressive parameter of an AR(1) model for annual outflows from the lake. The implicit moving average parameter θ depends on a dimensionless term σ_R^2 defined in equation (15) as the ratio of the variance of the annual net connecting channel flow to the variance of the annual change in lake storage. For certain values of these parameters ($0<\theta<\phi<1$ and ϕ close to 1), the ARMA(1,1) model leads to prominent tails in the ACF and, hence, mimics the effect of long-term persistence in the annual net basin supplies.

Except for the pairing between Lakes Superior and Michigan-Huron, both of which are moderated by Plan 77A, there is reasonable agreement between the sample and predicted CCFs. The CCF match is especially good among the net basin supplies to the small lower basin lakes. At the large upper basin lakes, both the sample and predicted ACFs drop to zero rapidly. At Lake Superior the autoregressive parameter is too small to sustain high lag autocorrelations. At Lake Michigan-Huron, the ACF does not reduce to an ARMA(1,1) form. Sample and predicted ACFs at the lower basin lakes exhibit a much more gradual decay. Here the autoregressive parameters are large enough to sustain a long tail ACF. The prominence of the tail in the predicted ACF depends on the variance ratio σ_R^2 which is found to increase as lake size decreases. As a consequence, the residual method leads to a covariance structure in the resulting net basin supplies that depends implicitly on lake scale.