Water Supplies to the Great Lakes—Reconstructed from Tree-Rings

W. A. R. Brinkmann

Department of Geography and Institute for Environmental Studies, University of Wisconsin, Madison, WI 53706

(Manuscript received 20 March 1986, in final form 13 November 1986)

ABSTRACT

Correlations between the water supplies to each of the Great Lakes and prewhitened tree-ring chronologies from 16 sites around the Great Lakes suggested some strong associations for the summer months, particularly June and July. Some of these associations were, however, nonlinear and variables had to be log-transformed. Regression models for June–July water supplies totals were developed for each lake. These models explained between 45 and 56% of the variance in the dependent datasets; they also performed satisfactorily on independent data sets. The models were therefore used to estimate past water supplies to each of the lakes as far back as the chronologies in the models permitted (about 200 years). The association between variations in supplies to the individual lakes during this century was compared to that of the reconstructed past supplies and was found to be similar. Variations in the water supplies to Lake Superior are to some degree independent of those to the lower lakes—which makes the use of Lake Superior as a storage reservoir in lake level regulations possible; there is, however, also a certain degree of similarity in supply variations between the lakes which, especially in the case of large supply anomalies, reduces the success of any type of regulation plan for the Great Lakes.

1. Introduction

Tree-rings have been used successfully in the reconstruction of such climate variables as temperature and precipitation. In addition, the reconstruction of hydrologic variables which, like tree growth, integrate the effects of temperature and precipitation, have been very successful. Examples of the latter are the reconstructions of stream flow in the Colorado and Gila Rivers (Stockton and Jacoby, 1976), water levels of Lake Athabasca (Stockton and Fritts, 1973), and Potomac River streamflow (Cook and Jacoby, 1983).

There is considerable value in the reconstruction of hydrologic variables because water use and regulation plans are usually based on short records which may not be representative of longer-term variations. The regulation of Lake Ontario, for instance, which was implemented in 1960, failed twice within the following decade and a half: during the early 1960s when water supplies to all lakes were extremely low, and again in the early 1970s when water supplies to all lakes were extremely high (IJC, 1976). A major problem with the regulation of the Great Lakes is the narrowness and shallowness of the connecting channels; these natural restrictions limit the speed with which water can be moved through the system. It takes, for instance, more than 2 yr for 50% of a change in water supplies to Lake Huron to be realized in the outflows from Lake Ontario (IJC, 1976).

Great Lakes regulations (currently in effect for Lake Superior and Lake Ontario) are designed to maintain lake levels within a certain range by regulating inflows and outflows. Such regulations are intended to reduce the occurrence of extremely high and low levels since these have detrimental effects on shipping, hydroelectric power generation, shore property, etc. Because of the low response time of the Great Lakes system (as a result of the natural restrictions), current lake level management makes use of the lake farthest upstream—Lake Superior—as a reservoir in which water is retained and stored at times of high supplies and from which water is released at times of low supplies.

The success of any regulation plan for a series of interconnected lakes, such as the Great Lakes, depends on the degree of intercorrelation between the supplies to the lakes, and between supplies to the reservoir lake and supplies to the lower lakes. Because of the low response time of the Great Lakes system, the higher the intercorrelations, the less successful any type of plan can be.

Previous studies have shown that water supplies to the Great Lakes respond more to variations in precipitation than to variations in temperature (Brinkmann, 1983); they have shown that a climate boundary (in terms of precipitation) divides the Great Lakes region into a northwestern and a southeastern portion (Brinkmann, 1983a); and they have shown that, therefore, variations in water supplies to Lake Superior are to some degree independent of variations in supplies to the lower lakes, but that very large supply anomalies can be of the same sign basin-wide and can persist for several seasons (Brinkmann, 1983b). Periods of basin-wide persistent anomalous supplies have led to the low lake levels of the 1930s and 1960s, and the high levels of the 1950s and 1970s. In the spring of 1986, following about 6 months of above normal precipitation, record
high lake levels existed on all of the Great Lakes, except Ontario. Record levels continued into the summer of 1986, even though below average precipitation fell in the spring.

The purpose of the present study was, therefore, to estimate variations in water supplies to the Great Lakes prior to the beginning of the observed record (using tree-ring data) and to determine whether the nature of the interrelationship between supplies to the lakes during this century has been similar to that of a longer past-time period.

2. Data

a. Lake data

Water supplies data rather than lake level data were used since levels incorporate past water supplies as well as inflows from upstream lakes, outflows to downstream lakes, and lake level regulations. Lake levels will, for instance, rise during a wet period; during a subsequent dry period, levels will initially still be high but falling. Furthermore, inflow and outflow are the largest components of the water budget for the lower Great Lakes and, consequently, determine their levels. The net basin water supplies, on the other hand, are those components of the water budget which are contributed by the lake basin itself: lake precipitation plus runoff minus lake evaporation. Long-term time series of net basin supplies to each of the Great Lakes cannot be computed directly from these variables since they cannot be measured directly for large water bodies. Net basin supplies to each of the lakes can, however, be derived from the available long records of lake level, flow, and diversion, using the water budget equation:

\[ \text{NBS} = dS + O - I \pm D \]

where NBS are the net basin supplies, \(dS\) is the change in water volume stored in the lake (which is computed from the change in lake level and area of the lake), \(O\) the outflow from the lake, \(I\) the inflow from the upstream lake, and \(D\) the total diversion. (An example of diversion of water into one of the Great Lakes is the diversion from the Albany River basin, through the Long Lake and Ogoki Projects in Canada, into Lake Superior; an example of diversion from one of the lakes is the water diverted out of Lake Michigan at Chicago.) The volume of ground water flowing into or out of any of the Great Lakes is unknown, but the net effect is believed to be negligible and is normally disregarded in Great Lakes water balance studies (Jones and Meredith, 1972).

Computed NBS show a seasonal cycle: NBS increase in spring because precipitation begins to increase and runoff is at its maximum; NBS are largest in late spring—early summer because precipitation is high and evaporation is low; NBS begin to decrease again in late summer and fall as precipitation decreases and evaporation reaches its maximum.

b. Tree-ring data

The tree-ring data consist of annual tree-ringchronologies from 16 sites around the Great Lakes. (A numbering system—C1 through C16—is used to indicate the site locations in Fig. 1 and to refer to the chronologies in the following discussion.) Eight of these chronologies (C1—C3, C11, C13—C16) were obtained from the archives of the Laboratory for Tree-Ring Research at the University of Arizona. Three chronologies (C4—C6) were provided by M. L. Parker and L. A. Joosza,1 and M. E. Alexander.2 The remaining five chronologies were collected by the author. All chronologies were collected and processed according to procedures described by Fritts (1976). These procedures include measuring and crossdating the annual rings, standardizing the ring widths by fitting a growth curve to each time series of rings widths and dividing each year’s width by the value of the fitted curve for that year. (Regarding the trees sampled by the author, two cores per tree were taken from opposite sides of the

---

1 Forintek Canada Corp., Western Forest Products Laboratory, Vancouver, British Columbia.
tree, and the rings were standardized by fitting a negative exponential curve or a straight line to the series.) The average over all standardized cores taken from a stand is the chronology for that stand. Each chronology is represented by a single tree species: red pine (C1–C11), hemlock (C12, C15), white pine (C14), pitch pine (C16), chestnut oak (C13). Based on the oldest tree in each chronology, the beginning years of the chronologies range from 1620 (C1) to 1773 (C9).

While the effect of climate for a year on tree growth is usually greatest during that year, climate may also affect physiological processes that influence growth in subsequent years. This results in high autocorrelations in tree-ring chronologies which tend to pose problems in the development of equations relating tree-ring width to climate or to hydrologic variables (called transfer functions). First-order autocorrelations in the 16 chronologies ranged from +0.26 to +0.74 for the period 1900–1971. The tree-ring chronologies were therefore prewhitened. Following the method used by Cook and Jacoby (1983), an autoregressive (AR) model of order $p$ (first, second, or third order) was fitted to each chronology. In contrast to the tree-rings, there is no significant persistence in monthly water supplies from one year to the next, and the NBS data were therefore not prewhitened.

3. Method

a. Transfer functions

Because previous work had indicated the presence of a climatic divide across the Great Lakes and a consequent difference between variations in NBS to Lake Superior and NBS to the lower lakes, transfer functions between water supplies and annual tree growth were developed separately for each lake.

For each lake, a regression model was developed using an “all-possible-subsets” procedure which examines all $2^p$ possible regressions (where $p$ is the number of predictors), and prints the “best” several equations for inspection by the investigator. The best model is selected on the basis of three criteria: the squared multiple correlation coefficient ($R^2$), the $R^2$ adjusted for the number of predictors in the model ($R^2_a$), and Mallow’s $C_p$ statistic (which trades off the goodness of fit of the model against the number of parameters in the model). This procedure is less sensitive to a variety of statistical assumption violations than a stepwise regression (Weisberg, 1980). While this procedure may result in spuriously high $R^2$, particularly when the number of observations is less than the number of predictors (Rencher and Pun, 1980), this is not estimated to be a significant problem in the present analysis since for each lake the number of observations is 60 and the number of candidate predictors is 16. The procedure has been successfully applied to the reconstruction of paleoclimatic data from fossil pollen (Bartlein et al., 1984) and from tree-rings (Graumlich and Brubaker, 1986).

Prior to the building of the models, potential problems related to nonlinearities were investigated, and the data were screened to determine for which time of year the association between tree-rings and NBS is strongest.

Examination of the simple correlations between monthly water supplies to the Great Lakes and the prewhitened annual tree growth showed that the association between them is strongest for the months of May–August. This time of year is approximately the radial growth season for trees in this region; it is also the time of largest water supplies to the lakes. For each lake, the highest correlations are for June and/or July and range from +0.35 to +0.54 (statistically significant beyond the 99% confidence level); but correlations between several chronologies and monthly water supplies for that time of year are small and some are negative. Since tree-ring width is a response to climate variations in the vicinity of the stand, it cannot be expected that all chronologies correlate highly with the water supplies to the lakes because the much larger lakes integrate climate over much larger areas. High correlations may, however, occur if the tree stand is relatively close to the lake and one of them is upwind of the other so that both are frequently under the influence of the same weather system (as, for instance, if both are located in the path of a prevailing storm track). Positive correlations between a lake and distant tree stands may reflect climatic homogeneity within the Great Lakes basin. Negative associations, on the other hand, may reflect climatic differences within the Great Lakes basin.

Nonlinearities in the association between water supplies to the lakes and tree-ring widths were considered a possibility, based on theoretical considerations as well as on the work by, for instance, Cook and Jacoby (1983) and Duvick and Blasing (1981). Duvick and Blasing suggested that a nonlinear relationship between tree-ring widths and precipitation in central Iowa may be due to a weaker association between them at times of high precipitation totals when growth is limited by factors unrelated to precipitation; they were unable to improve their reconstruction through transformation of the variables and concluded that all that can be said about years corresponding to very wide rings is that they were wet. Cook and Jacoby (1983), in their reconstruction of Potomac River streamflow, log-transformed the streamflow series.

In the present study, the possibility of nonlinear relationships was therefore recognized, and the relationships were investigated by constructing bivariate scatter diagrams. (The exact nature of the association between an individual tree stand and the water supplies to one of the Great Lakes is, however, more complex—as discussed above—that between a stand and the precipitation record from a nearby weather station.) Inspection of the scatter diagrams revealed nonlinearities
in the relationship between some tree-ring chronologies and some lakes, and the form of the scatter for those cases suggests that ring-width increases as NBS increases, but the increase in width levels off at high NBS. Because of the general lack of published documentation of nonlinear associations, an example of a scatter diagram is presented in Fig. 2. It is a plot of June–July NBS to Lake Superior against prewhitened tree-rings from a stand located just to the northwest of Lake Superior (C1). Both the lake and the stand are located in the path of one of the primary summer cyclone tracks across North America: Cyclones generated over Wyoming affect the northwestern portion of the Great Lakes basin before heading northeastward to James Bay (Whittaker and Horn, 1984). The scatter in Fig. 2 shows not only a leveling off in ring width with increasing NBS but there is also some indication of a decrease in width with very high NBS (which could mean that while there is no limit to the amount of water received by the lake, growth conditions may become less optimum when moisture supplies are very high). The best linearization of the relationship shown in Fig. 2 was obtained through log-transformation of the NBS as well as the ring widths. Inspection of all scatter diagrams showed that for the upper two lakes (Superior and Michigan/Huron) the strongest relationships between NBS and individual chronologies (in terms of the correlation coefficient) were similar to the one shown in Fig. 2. Consequently, the NBS for these two lakes and those important nonlinearly related chronologies were log-transformed. Some of the relationships between the lower two lakes (Erie and Ontario) and individual chronologies also appeared to be nonlinear but transformations of the variables were not found to be effective.

The data were also checked for inconsistent values, or outliers, using the studentized residuals, Mahalanobis distance, and Cook’s distance (Weisberg, 1980). Outliers may, for instance, be caused by insect infestation or wind damage. Outliers may also be caused by the occurrence of, for instance, a dry spring followed by a wet summer; since the trees would integrate the effects of both seasons, the reconstructed summer moisture conditions for that year would be poor. (This effect caused Stahle et al., 1985, to eliminate three data points.) In the present analysis, between two and three observations per lake had to be eliminated.

b. Verification procedures

To test the performance of the regression models, the 72 years for which both NBS and tree-ring data overlap (1900–1971) were divided into a calibration set consisting of 60 years and an independent or verification set consisting of 12 years. The 12 independent years were selected using a method that attempts to ensure that the verification set has a distribution of the dependent variable, NBS, similar to the calibration set and is at the same time based on a random selection procedure. The 72 NBS values for each lake and summer month (and combinations of summer months) were rank-ordered and then divided into six subsets, each consisting of 12 years. From each of the six subsets, two years were chosen, using random numbers, to be included in the verification set. This approach not only assures a similar distribution of the dependent variable but also reduces the problem resulting from the inclusion of a sequence of unusual years in the independent set, such as the “bridging effect” described by Cook and Jacoby (1983) which caused them to eliminate several data points from their independent set.

Two verification measures were used to test the performance of the regression models on the independent data: the square of the correlation coefficient ($R^2$) and the reduction of error statistic (RE). The RE is a very rigorous and sensitive (to even a small number of bad estimates) measure for which, however, no formal significance test exists (Gordon and Le Duc, 1981). In general, RE is equal to 1 if the agreement between prediction and observation is perfect; a positive value of RE is an indication of some skill in the sense that the predictions are better than simply predicting the average of the calibration set.

For Lake Superior, the model was verified on an additional 40 years (1860–1899). Because this series of 40 years was not selected in the same manner in which the 12-yr independent sets were selected, the regression model for Lake Superior was tested on not only this series of 40 years but also on a reduced dataset (with two very unusual data points eliminated from the 40 years) and on a smoothed dataset (applying a three-term—$1/4, 1/2, 1/4$—binomial filter to the 40 years).

Since the NBS are reconstructed separately for each lake, empirical orthogonal function (EOF) analysis was used to test the similarity in spatial patterns between the observed and reconstructed NBS to the four lakes. This type of analysis consists of extracting from a
of the loadings, one for each lake. (Loadings are usually plotted on a map but since there are only four data points in the present analysis—one for each lake—the loadings are more conveniently shown in the form of bars.) Loadings that are of the same sign indicate similarity in the variation of the variable—or NBS to the

number of maps a small number of patterns or EOFs which explain a large fraction of the total variance of the observed field and thus represent the dominant modes of spatial variation of the field with time. The particular method used in the present study is the one described by Kutzbach (1970).

4. Results

Models for each summer month and combination of summer months (with the 16 tree-ring chronologies as the candidate predictors) were investigated. As had already been suggested by the simple correlation coefficients, the best models were for June–July NBS totals. For each lake, the model consists of seven predictors (Table 1). The unadjusted variance explained by the calibration models ranges from 45 to 56%. Verification of the models on the independent data produced significant correlation coefficients and positive RE statistics (Table 2).

To compare the spatial patterns of the observed and the reconstructed June–July NBS for the period 1900–1971, EOFs were computed. For both the observed and the reconstructed NBS, the first two EOFs together account for over 80% of the total variance and thus represent the most dominant modes of variation (Fig. 3). The bars in Fig. 3 represent the magnitude and sign

---

**Table 2. Calibration and verification statistics.**

<table>
<thead>
<tr>
<th>Predictand</th>
<th>Test on dependent dataset</th>
<th>Test on independent dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>$R^2_a$</td>
</tr>
<tr>
<td>Superior</td>
<td>.54</td>
<td>.47</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michigan–Huron</td>
<td>.56</td>
<td>.50</td>
</tr>
<tr>
<td>Erie</td>
<td>.45</td>
<td>.40</td>
</tr>
<tr>
<td>Ontario</td>
<td>.51</td>
<td>.44</td>
</tr>
</tbody>
</table>

* Significant at the .05 level.
** Significant at the .01 level.

---

**Fig. 3.** Loading pattern of the first EOF (top) and second EOF (bottom) of the observed net basin supplies (solid line) and reconstructed supplies (dashed line) for June–July, 1900–1971. The row of numbers, labeled EV, gives percent of variance for each lake accounted for by the EOF; values for the reconstructed supplies are given in brackets.
lakes; loadings of opposite signs indicate out-of-phase relationships. The loadings for the first EOF of the observed NBS are of the same sign for all four lakes but the magnitude is smallest for Lake Superior. The first EOF of the observed NBS accounts for 59% of the total variance; it accounts, however, for 58% to 80% of the variance in the supplies to the lower three lakes and for only 28% of the variance in the supplies to Lake Superior. The spatial pattern of both the loadings and the variance accounted for by the first EOF thus reflect the degree of similarity in supply variations between the lakes and reflect the degree of climatic homogeneity within the Great Lakes basin, which is largest across the southeastern portion of the basin and the lower three lakes. For the second EOF of the observed NBS, the loading is large and positive for Lake Superior; for the lower three lakes, the loadings are smaller and variable in sign. The second EOF accounts for 23% of the total variance; it accounts, however, for 67% of the variance in the supplies to Lake Superior and for only 1% to 20% of the variance in supplies to the lower lakes. The spatial pattern of the loadings and the variance accounted for by the second EOF thus reflects the degree to which the supplies to Lake Superior are independent of the supplies to the other lakes and thus the degree of climatic nonhomogeneity across the Great Lakes basin. The EOFs of the reconstructed dataset for the period 1900–1971 (Fig. 3) are very similar to those of the observed NBS, which show that although the NBS were reconstructed separately for each lake, the reconstructions accurately portray the associations between the lakes.

Given the relatively sparse, but wide, spatial coverage of tree-ring sites, the models performed well, and reconstruction of past variations in NBS to the lakes was considered feasible. Estimates of past June–July NBS to the four lakes were therefore computed from the tree-ring data, going as far back in time as the chronologies permitted (Fig. 4). The length of the reconstructed time series varies from lake to lake since it is a function of the shortest tree-ring chronology in the regression model for the lake. The first year of a chronology is often based on only a single tree core, which is a very small sample for reconstruction. However, within 7 years of the beginning of the shortest chronology in each lake model, the sample size increases to more than five.

5. Discussion and conclusion

The time series of reconstructed NBS show a tendency for above-average values during the first half of the 1800s and below average values during the second half and the latter part of the 1700s. Comparison between these reconstructions and some of the historical lake level records (Foster and Whitney, 1851; Horton, 1927; IJC, 1976) is difficult since lake levels integrate over a sequence of seasons and years (lake levels will, for instance, be highest at the end of a period of positive NBS anomalies). However, reports of high lake levels on some of the lower lakes during the late 1830s, for instance, do correspond with the end of a period of reconstructed high NBS to the Great Lakes, particularly the lower lakes. Furthermore, climate data for the Great Lakes region indicate that the summers of the 1830s were wet over almost the entire Great Lakes basin except for the western portion of Lake Superior (Wahl, 1968). Such data also indicate that the summers of the 1850s and 1860s were dry over the northwestern portion of the basin while near-average conditions prevailed over the rest of the region (Wahl and Lawson, 1970), which is also in very good agreement with the reconstructed low supplies to the upper lakes and near average supplies to the lower lakes.

Of considerable importance, for the purpose of the present study, is whether the degree of intercorrelation between the net basin supplies to the four lakes that has existed during this century (Fig. 3) is similar to that of a longer past time period. Results of previous EOF analysis of observed NBS patterns for all calendar months (Brinkmann, 1983b) are similar to those for June–July shown in Fig. 3; there is presently a large positive intercorrelation between the lakes (as indicated by the first EOF) which represents a severe problem to the regulation of the Great Lakes system. Variations in the supplies to Lake Superior are, however, to some degree independent of the supplies to the lower lakes (as indicated by the second EOF). It is this low correlation between Lake Superior and the lower lakes—as a result of a fortunate precipitation divide across the basin—that makes the use of Lake Superior as a storage reservoir possible. Extreme supply anomalies can, however, be basin-wide and lead to extreme levels.

The EOF analysis of the long record of reconstructed NBS to the four lakes suggests that the past patterns of intercorrelation were very similar to those of the present century (Fig. 5); there has been a considerable degree of intercorrelation between the lakes, but with some independence between the supplies to Lake Superior and the supplies to the lower lakes. Independence means, of course, that the supplies to Lake Superior and the lower lakes are sometimes of the same sign and sometimes of opposing signs. Another important question therefore, considering the slow response time of the Great Lakes system, is the degree of intercorrelation for extreme supply conditions. A tabulation was made of the summers (June–July) for which the reconstructed (1773–1899) supplies were extreme (defined for the present purpose as exceeding ±1.00 s.d.). For about one-quarter of the summers for which the supplies to Lake Superior were extremely high, two or more of the other lakes also received extremely high supplies. For almost one-half of the summers for which the supplies to Lake Superior were extremely low, two or more of the other lakes also received extremely low supplies. These frequencies are not only very similar
Fig. 4. The reconstructed (solid line) and observed (dashed line) June–July net basin supplies to the four lakes. The series are smoothed using a three-term binomial filter (1/4, 1/2, 1/4).
to those for the observed supplies for this century, but they also show that the reconstructed data accurately portray the fact that basin-wide high supplies are less frequent than basin-wide droughts because of the spatial scale and predominant tracks of precipitation producing systems: convective precipitation produced by individual thunderstorms is relatively frequent in summer and is very localized; frontal precipitation is more widespread but cyclones in summer tend to track either across the northwestern or the southeastern portion of the basin (Whittaker and Horn, 1984).

Another feature of interest is the episodic character of net basin supplies. Departures from average supplies are not randomly distributed in time but show a strong tendency for persistence (which must be due primarily to the character of long-term climate fluctuations). Thus, many of the summers (June–July) for which the reconstructed supplies to Lake Superior and to at least two other lakes exceeded +1.00 s.d. occurred during the first half of the 1800s. Almost all of the summers for which the reconstructed supplies to Superior and at least two other lakes were less than −1.00 s.d. occurred during the latter part of the 1700s and the second half of the 1800s.

Acknowledgments. I thank L. Graumlich, A. Swain, M. Winkler, and K. Gijewski for help with the collection and analysis of the tree-ring data. This work was funded by the University of Wisconsin Sea Grant College Program under a grant from the Office of Sea Grant, National Oceanic and Atmospheric Administration, U.S. Department of Commerce and by the State of Wisconsin.

REFERENCES


