NOAA Technical Report NOS CS 1

THE ACCURACY OF GREAT LAKES MEAN WATER LEVEL COMPUTATIONS USING REDUCED NETWORK CONFIGURATIONS

Chris Zervas
August 1997

noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT
OF COMMERCE
William Daley, Secretary

National Oceanic and
Atmospheric Administration
D. James Baker, Under Secretary

National Ocean Service
Nancy Foster
Assistant Administrator

Office of Coast Survey
Frank Maloney

Coast Survey Development
Laboratory
Bruce Parker

NOTICE

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The National Ocean Service (NOS) has been responsible for operating and maintaining the U. S. water level gauges on the Great Lakes and interconnecting waterways since NOAA's inception in 1970. Previously, the gauges were installed and operated by the U. S. Army Corps of Engineers who are charged with regulating water levels on the lakes in cooperation with the Canadian Hydrographic Service. The outflows from Lakes Superior and Ontario are controlled by the International Joint Commission, a committee formed by Canada and the U. S. (Grima and Wilson-Hodges, 1977). All the lakes are affected by diversions and modifications to the waterways between lakes (Hartmann, 1988; David et al., 1988). Effective regulation of the lakes requires timely and accurate measurement of the water level in each lake.

The Army Corps of Engineers (Detroit District) uses the Great Lakes water level data to calculate monthly averages for each lake and publishes the lake levels in the "Monthly Bulletin of Lake Levels for the Great Lakes". The publication also gives a six-month lake level forecast based on future weather conditions (Clites, 1992). The Canadian Hydrographic Services publishes a similiar bulletin. The monthly averages are determined using six stations each for Lake Ontario and the Lake Michigan-Lake Huron system, five stations for Lake Superior, four stations for Lake Erie and two stations for Lake St. Clair. The main reason for averaging over several stations is to minimize the effect of the long-term tectonic uplift of the northern lake shores relative to the southern lake shores due to glacial rebound (Tovell, 1979; Clites, 1992). The water level network was recently releveled in 1985 to account for vertical displacements that had taken place since the previous datums were established in 1955 (Coordinating Committee, 1995).

For regulation of the lake levels, the Corps of Engineers requires a more current lake level value than the monthly value. They use the beginning-of-month lake level, which they obtain by averaging the daily values from the last day of the previous month and the first day of the current month (Quinn and Todd, 1974). This 2-day average is subject to wind set-up errors which can occur when a strong wind stress exists along or across the axis of the lake, tilting the lake's surface (Hamblin, 1987). Therefore to minimize this error, the water levels from a number of stations around the lake are combined using a method called the Theissen polygon method. The beginning-of-month lake levels are then used to obtain the rates of change of lake storage (Quinn and Todd, 1974). These statistics are useful for monitoring lake hydrology, shoreline erosion, navigation, and hydroelectric power generation. They are also used in water budget calculations of evaporation rates and groundwater influx.

This study demonstrates a method for determining the effect of a reduced network size on mean lake level accuracy. In the next section, four different methods of calculating mean lake levels using water level time series at multiple locations are described. Then, the present network configuration and the data sets to be used are introduced. In the subsequent section, 2-day averaged water level data are analyzed to examine the calculation of beginning-of-month mean lake levels. Later, an analysis is carried out with hourly data to evaluate the accuracy of calculating mean lake levels and lake surface tilt on an hourly basis. Finally, the results of these analyses are used in discussing how accurately the hourly water level time series at any individual station may be reproduced by a combination of other stations on the lake.

follow through
53-end

2. FOUR METHODS FOR ESTIMATION OF MEAN LAKE LEVELS

2.1. AVERAGE

The simplest method of calculating the mean lake level is to add the water level at all available stations together and then divide by the number of stations. This uniform weighting method will work well if there are enough stations available so that any other signals present (at one or more stations) cancel out exactly or are substantially diminished in amplitude. However, during periods of strong wind stress, tilting of the lake surface may cause errors in mean lake levels calculated by simple averaging.

2.2. THEISSEN POLYGON METHOD

This method of combining water level data from the periphery of a lake to obtain the mean lake level was described in a series of papers published by the Great Lakes Environmental Research Laboratory (GLERL) in the mid-1970s (Quinn and Todd, 1974; Quinn, 1975a, 1975b; Quinn and Derecki, 1976a, 1976b). The method is a weighted average with the weight for each station given by the Theissen polygon procedure. This procedure was developed by hydrologists to determine the mean precipitation in a basin based on measured precipitation at a limited number of stations (Croley and Hartmann, 1985). Polygons are drawn around each station with the edges of the polygons bisecting lines drawn between each pair of stations. The fractional area of each polygon relative to the area of the whole basin is the weight assigned to the station at the center of the polygon. The weighting is based solely on the geometry of the network and not on the signals recorded at each station.

This method was applied to the Great Lakes water level measurements where, unlike with precipitation data, the stations are all located along the edges of the basin. As a result of the elongated geometry of the lakes, stations near the middle of the lakes are more heavily weighted than stations at the ends of the lakes. The GLERL reports obtained Theissen weights for each new network formed as a new station was added to the existing stations over time and compared the differences in mean lake levels resulting from the addition of one station. When the differences became small, the network was judged to be adequate for measuring mean lake levels. Since the reports were published, two water level stations (Barcelona on Lake Erie and Two Harbors on Lake Superior) have been removed. For the Theissen weights to be used in this paper, the weight for Barcelona has been combined with that of the station at Erie and the weight for Two Harbors has been combined with that of the station at Duluth.

2.3. CROLEY'S SPATIAL-OPTIMUM METHOD

In the mid-1980s, Croley (1986, 1987) investigated methods of calculating weighted averages of Lake Erie and Lake Superior station data to eliminate long-term (weekly to monthly) wind set-up error. The theoretical wind set-up for a steady-state wind stress of unit amplitude was calculated for all the stations around the lake based on a numerical hydrodynamic model. Theissen weights were then calculated for every possible subnetwork composed of subsets of the complete network. The

errors in the Theissen mean lake level due to wind set-up were obtained for each subnetwork and minumum error networks were found. Errors were further reduced when weights were obtained without constraining them to be Theissen weights but subject to eliminating the long-term wind set-up error and minimizing the total error for twelve years of daily data. Croley called this the spatial-optimum method.

For smaller networks, the station weights were nearly uniform. However, for the larger networks, the southern shore of Lake Erie was more heavily weighted than the northern shore and the northern shore of Lake Superior was more heavily weighted than the southern shore. Using these Croley spatial-optimum weights to calculate mean lake level could cause errors over the long term due to differential glacial rebound (Clites, 1992). As mentioned in the previous section, for the calculations to be made in this paper, the station weight for Barcelona was combined with Erie and the station weight for Two Harbors was combined with Duluth. In addition, the station weight for Monroe on Lake Erie was combined with the weight for the station at Fermi.

2.4. EMPIRICAL ORTHOGONAL FUNCTIONS

When water levels are measured around the circumference of a lake, the result is a number of non-orthogonal time series. Several signals caused by different physical phenomena are combined in different proportions to form the total signal at each station. The mean lake level is a signal that should be present at each station with equal amplitude. The wind set-up signal (lake surface tilt) should also be present at each station but the amplitude will vary from station to station. The signal will be large in the upwind direction and large (but 180 degrees out of phase) in the downwind direction. The signal will have small amplitudes approaching a nodal line near the middle of the lake where the amplitude goes to zero. Other signals (possibly wind-driven) may be large at one or two of the stations and negligible at the other stations.

The empirical orthogonal function (EOF) method is a way of resolving independent, orthogonal signals from a number of non-orthogonal time series (Kundu et al., 1975; Preisendorfer, 1988). This is done by forming a symmetric matrix composed of the cross-correlations $R(z_i,z_j)$ of each time series $v_k(z_i)$ at station z_i with every other time series at station z_j. Auto-correlations $R(z_i,z_i)$ are along the diagonal of the matrix.

$$R(z_i,z_j) = \frac{1}{K} \sum_{k=1}^{K} v_k(z_i)\ v_k(z_j) \tag{1}$$

where k is an index for time and i and j are indices for the stations. There are K data points and N stations. This matrix is used to solve for the eigenvalues λ_n and eigenvectors $\phi_n(z_i)$ of the orthogonal signals or modes.

$$\sum_{i=1}^{N} R(z_i,z_j)\ \phi_n(z_i) = \lambda_n\ \phi_n(z_j) \tag{2}$$

where n is an index for the mode. The number of modes will be equal to the number of stations N. The eigenvalues indicate the variance or energy for each mode in the system. The eigenvectors indicate the amplitude or scaling factor for each mode at each station. A time series E_{kn} for each mode is also obtained which is a combination of the input time series at each station.

$$E_{kn} = \sum_{i=1}^{N} v_k(z_i)\ \phi_n(z_i) \tag{3}$$

This method is based on the actual signals recorded at the stations rather than on the geometry of the network. If the empirical orthogonal function analysis is carried out for the complete network and the first mode has nearly the same eigenvector amplitude at each station, the first mode is the mean lake level. It is implicit in this assumption that there are presently enough stations on each lake to closely approximate the mean lake level. Once the mean lake level time series is established, any other time series computed from fewer stations can be statistically evaluated to show how closely it approximates the mean lake level. The two statistics to be considered are the standard error and the maximum error.

3. DATA ASSESSMENT

3.1. DATA

There are presently 52 water level stations on the Great Lakes (NOS, 1994) without including stations on the waterways connecting the lakes. There are 31 U. S. stations and 21 Canadian stations (Figure 1). Archived hourly data for 1990 to 1994 were obtained for all 52 stations. The water level at each station is given to the nearest centimeter. The station at Mackinaw City is located on the Straits of Mackinac which connects Lake Michigan to Lake Huron. Since water can be transported in either direction through the straits, the Mackinaw City station is considered to be part of both the Lake Michigan and the Lake Huron networks.

All of the analyses in this report were carried out for two data sets. An analysis was carried out for 5 years (1990-1994) of 2-day averaged data to evaluate station networks necessary for obtaining the beginning-of-month lake levels used for lake level regulation. Although the beginning-of-month lake levels are dependent on only 2 days of data in a month, mean lake levels can also be calculated for any other 2-day period in a month. All 2-day periods in the 5-year data set were used in the analysis to provide an adequate sampling of high wind stress events that are more likely to occur during the winter months. Further analysis was also carried out with 1990 hourly data to examine the consequences of reduced network size on measuring both mean water levels and lake surface tilt on an hourly basis.

The hourly time series for 1990 for each lake are shown in Figures 2 to 6. Time series are offset to display all the stations. There are fourteen stations on Lake Erie, twelve stations on Lake Huron, and nine stations each on Lakes Superior, Michigan, and Ontario. The 2-day averaged time series for 1990-1994 are shown in Figures 7 to 11. Again, the time series are offset for comparison. Whenever any gaps in the hourly or 2-day averaged data occured at any station, all the data for the other stations on the same lake during the gap were dropped.

3.2. PRELIMINARY ANALYSES

Preliminary EOF analyses of Lake Michigan and Lake Huron data showed that in each lake one station was dominating the second modes; Green Bay for Lake Michigan and Essexville for Lake Huron. When one station dominates the second mode, it indicates that the second mode is not the general wind set-up over the whole lake but rather a large signal that is unique to that station. Each of these stations is at the head of a shallow bay that is at some distance from the main part of the lakes. The signals at these stations have larger amplitudes than the other stations, probably due to large wind set-up in the bays and in the case of Green Bay a resonance or seiche amplified by the bay (Figures 3 and 4). In order that the EOF analyses better represent the lakes as a whole, these two stations were eliminated from subsequent analyses.

Preliminary EOF analysis of the 14 Lake Erie stations showed the effect of strong wind set-up events near the two ends of the lake. The first mode turned out to be a combination of the mean lake level and the wind set-up at stations near the western end of the lake while the second mode was dominated by wind set-up at stations near the eastern end of the lake. This is due to the shallow depth of the western end of the lake which is very responsive to wind events. Only when two of the five western stations are dropped from the EOF analysis does the first mode represent the mean lake level alone and the second mode represent the wind set-up at both ends of the lake. Therefore, the stations at Toledo and Fermi are dropped and the twelve remaining stations are analyzed as the full network. However, the stations at Toledo and Fermi will be considered for possible smaller network configurations.

Figure 1. Great Lakes water level stations. There are 31 U.S. stations and 21 Canadian stations. (PC -- Port Colborne, PW -- Port Weller)

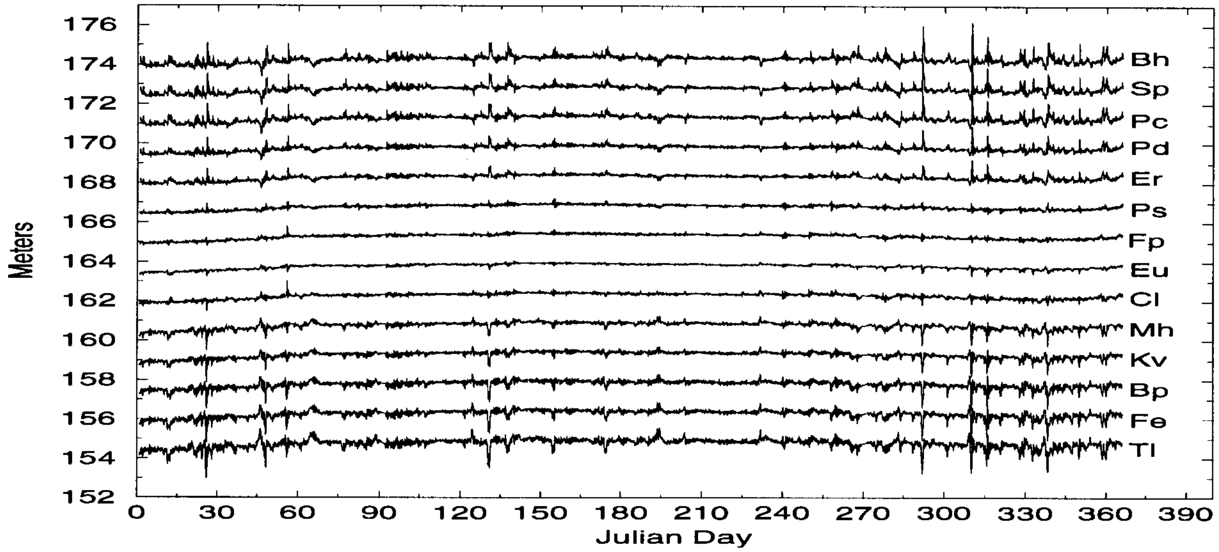


Figure 2. Lake Erie hourly water levels for 1990 (offset for comparison). Bh - Buffalo Harbor, Sp - Sturgeon Point, Pc - Port Colborne, Pd - Port Dover, Er - Erie, Ps - Port Stanley, Fp - Fairport, Eu - Erieau, Cl - Cleveland, Mh - Marblehead, Kv - Kingsville, Bp - Bar Point, Fe - Fermi, Tl - Toledo.

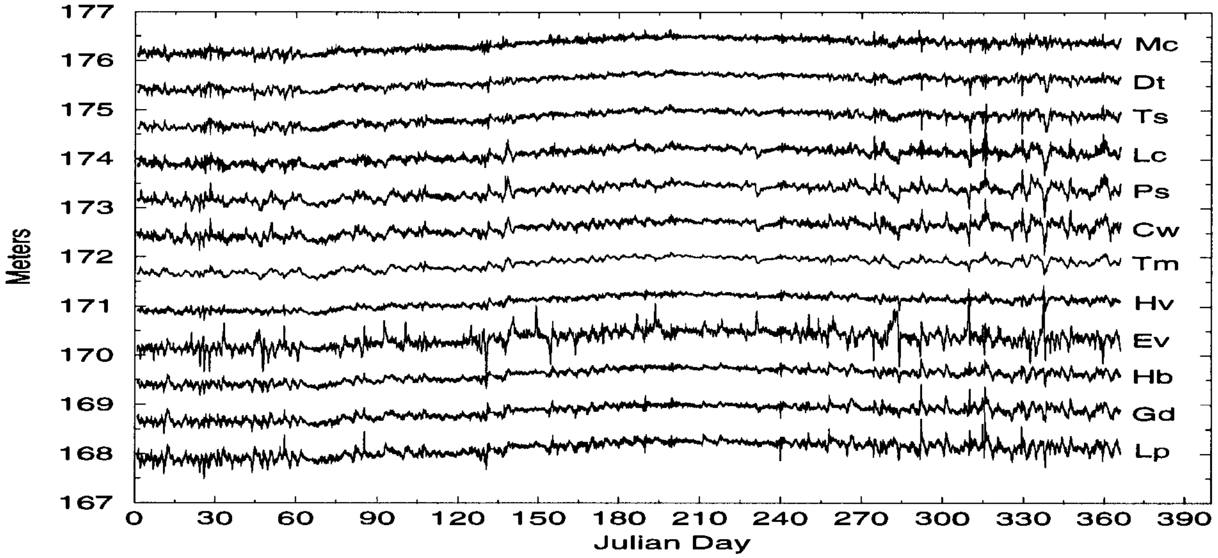
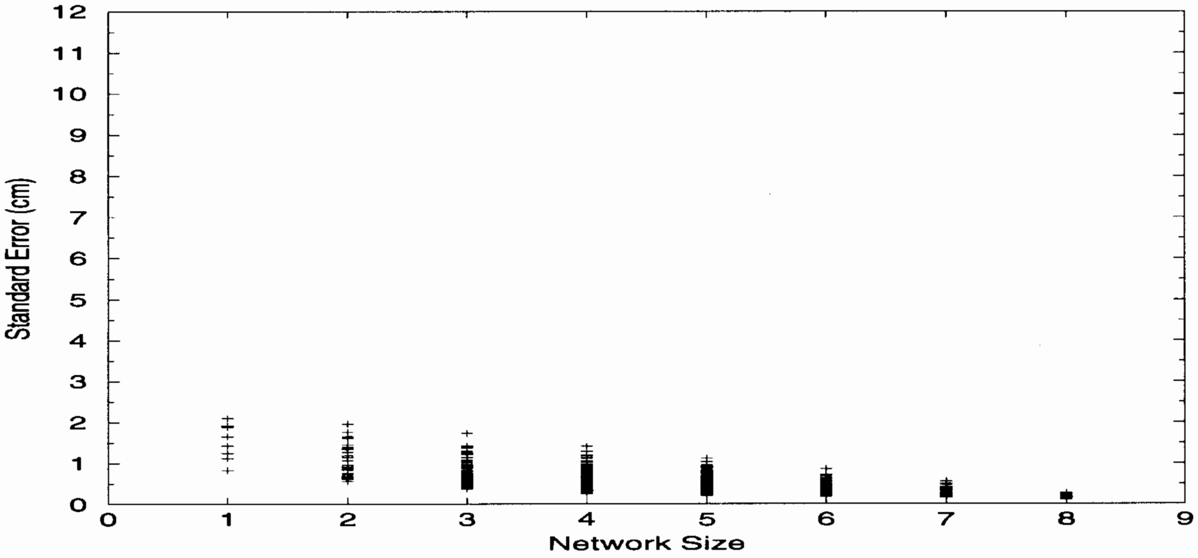
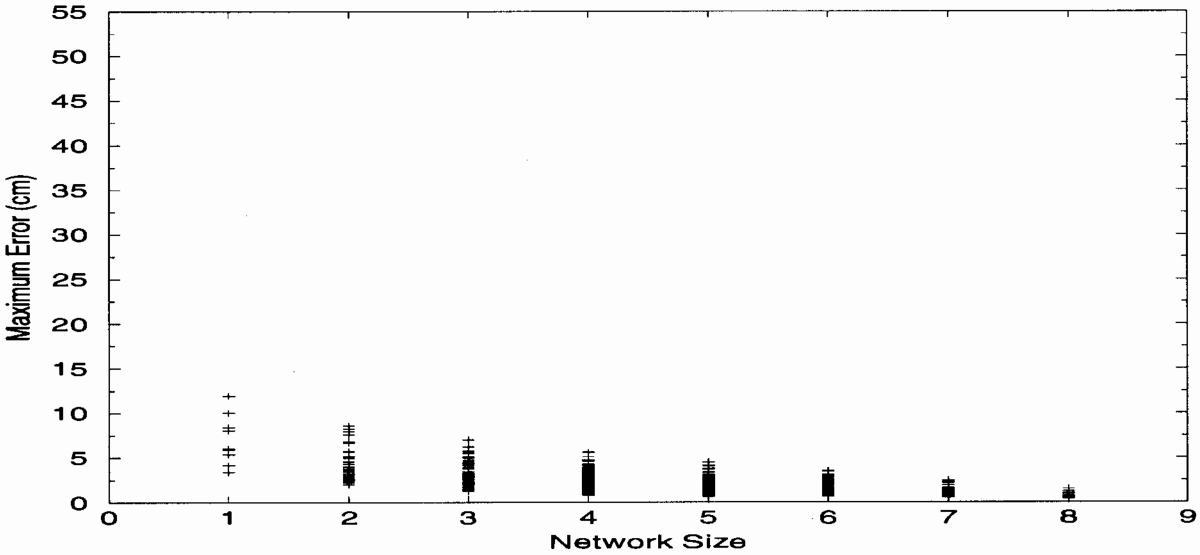


Figure 3. Lake Huron hourly water levels for 1990 (offset for comparison). Mc - Mackinaw City, Dt - De Tour, Ts - Thessalon, Lc - Little Current, Ps - Parry Sound, Cw - Collingwood, Tm - Tobermory, Hv - Harrisville, Ev - Essexville, Hb - Harbor Beach, Gd - Goderich, Lp - Lakeport.

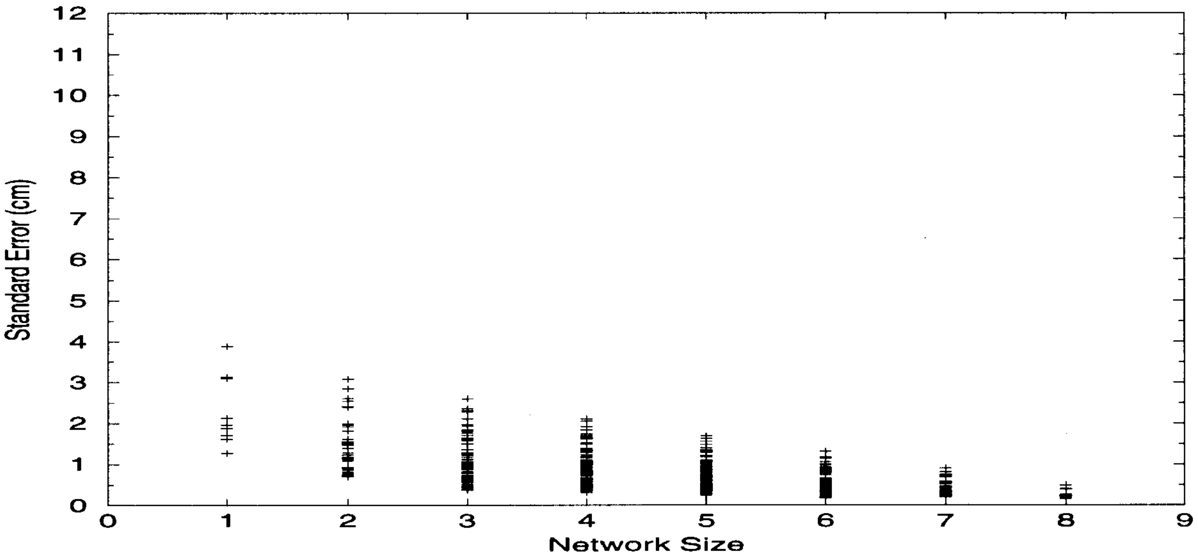


a)

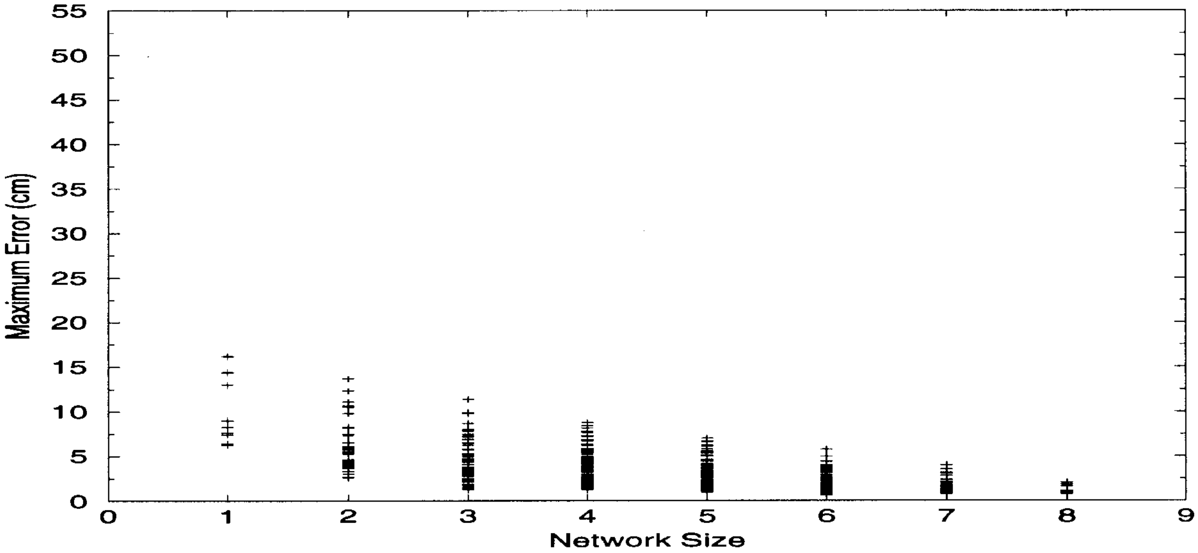


b)

Figure 20. a) Standard error and b) maximum error of subnetwork first EOF modes for Lake Ontario 2-day averaged lake levels.

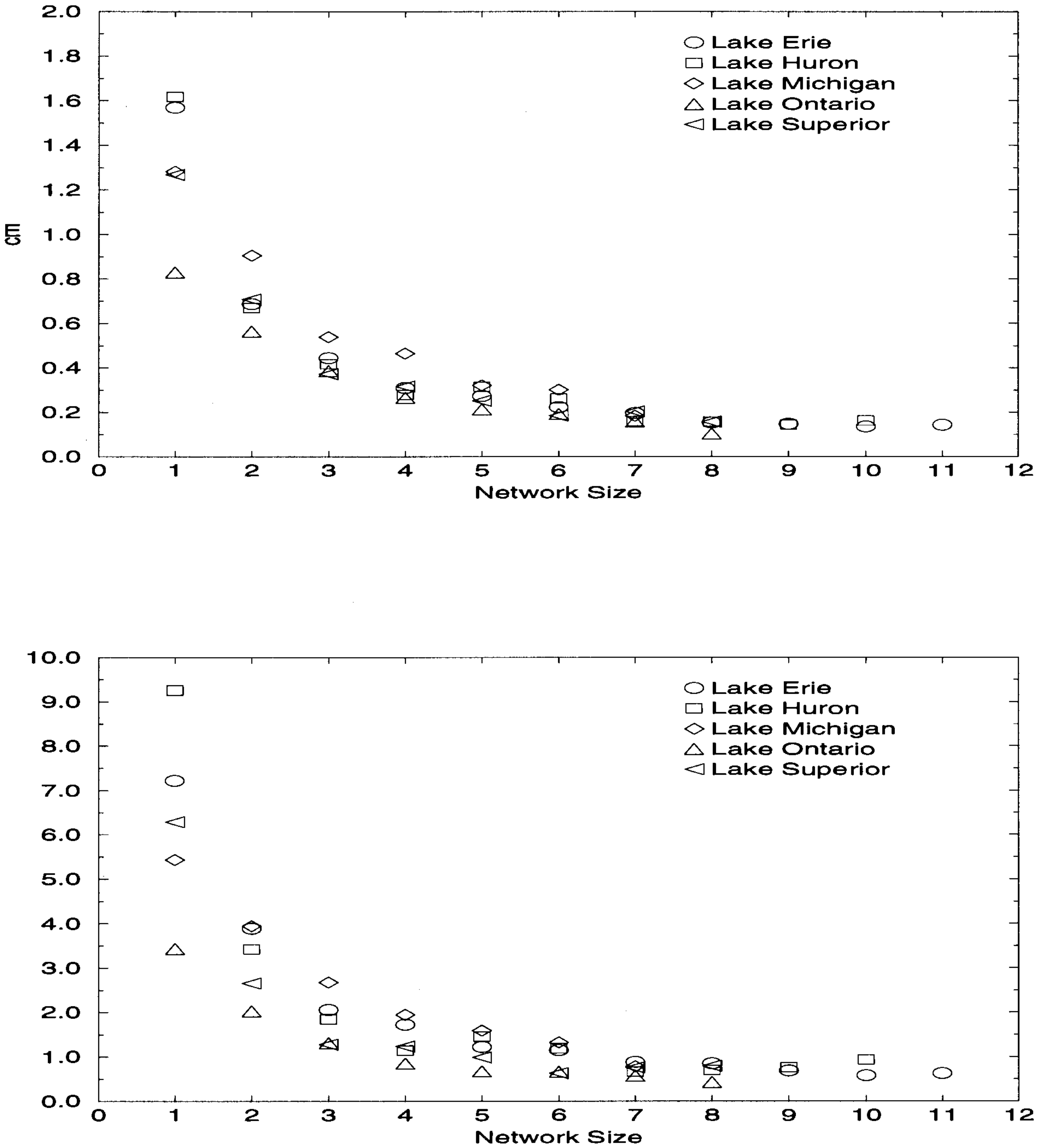


a)



b)

Figure 21. a) Standard error and b) maximum error of subnetwork first EOF modes for Lake Superior 2-day averaged lake levels.



a)

b)

Figure 22. Lowest first EOF mode a) standard error and b) maximum error for 2-day averaged lake levels.

5. ANALYSIS OF HOURLY DATA

5.1. MEAN LAKE LEVEL COMPARISONS

The four different methods of calculating mean lake levels from hourly data for 1990 are shown in Figures 23 to 27 (offset for comparison). The first EOF mode can be compared to the average, Theissen, and Croley (for Lakes Erie and Superior) methods. The second EOF modes representing lake surface tilt along the longer lake axis (plus the third EOF mode for Lake Huron) are also shown for comparison with the first EOF modes. All mean lake levels produce similar annual cycles with some high frequency differences during storms. As with the 2-day averaged data analyses in the previous section, Lakes Huron and Michigan mean lake levels have more high frequency variability. As before, the Theissen method gives a slightly smoother curve since stations near the middle of the lake are more heavily weighted than stations near the ends of the lake. If the Theissen method is assumed to produce the closest approximation to the mean lake level, then the resulting difference statistics relative to the Theissen method (standard error and maximum error) are shown in Table 3. Standard errors are less than 2 cm for all methods. Although none of the methods clearly stands out as the best overall, the EOF method has other desirable properties.

**Table 3. Mean Lake Level Difference Statistics (cm)
relative to the Theissen Method
1990 hourly data**

	Method	Standard error	Maximum error
Lake Erie	Average	1.71	16.60
	Croley	1.48	15.00
	EOF	1.91	8.94
Lake Huron	Average	0.99	7.60
	EOF	1.60	11.95
Lake Michigan	Average	0.91	6.00
	EOF	1.21	5.02
Lake Ontario	Average	0.57	4.00
	EOF	0.70	3.44
Lake Superior	Average	0.75	4.50
	Croley	0.80	4.60
	EOF	1.01	4.04

5.2. MODAL ENERGY DISTRIBUTION

The EOF mode energy percentages for 1990 hourly data from the complete network for each lake are shown in Table 4. Compared to the 2-day averaged data, the percentage of energy is lower for the first mode and greater for the second mode. This is because the time averaging of the hourly data in the previous analyses reduced the amplitude of the lake surface tilt signal. The percentage of energy in the first mode ranges from 61.7% for Lake Erie to 98.0% for Lake Ontario. The amount of energy in the second mode, which is due to lake surface tilt along the longer axis of the lakes, ranges from 34.0% for Lake Erie to 1.1% for Lake Ontario. Lake surface tilt along the shorter axis is the third mode for Lakes Superior, Huron, and Ontario and the fourth mode for Lakes Michigan and Erie. The third mode of Lakes Michigan and Erie has the middle of the lake out of phase with the two ends. The third mode is comparable in energy to the second mode only for Lake Huron. These results are similar to the results obtained with 2-day averaged data.

Table 4. EOF Mode Energy Percentage (Hourly Data)

Mode	1	2	3	≥4
Superior	85.9	9.9	1.7	2.5
Michigan	87.7	7.4	1.4	3.5
Huron	87.6	6.0	3.7	2.7
Erie	61.7	34.0	1.6	2.7
Ontario	98.0	1.1	0.3	0.6

5.3. NETWORK SIZE

The consequences of using smaller network sizes to obtain both the first EOF mode and the second EOF mode were evaluated using hourly data for 1990. (In the previous section, using 2-day averaged data, only the first EOF mode errors were evaluated.) The error statistics in this case were calculated for the difference between the first mode of the complete network and of the subnetwork plus the difference between the second mode of the complete network and of the subnetwork. If the first two modes can be reproduced by a subnetwork, most of the water level variability in the lake is being measured. The error statistics are shown in Figures 28 to 32 for every possible subnetwork as a function of network size.

The errors for hourly data are greater than the errors for 2-day averaged data, since the lake surface tilt signal is stronger in the hourly data and now the error statistics for the first two EOF modes are being considered. Lake Erie has large ranges of errors for each network size due to its responsiveness to wind stress. Subnetworks with all stations near one end of the lake can have large errors (standard errors greater than 10 cm and maximum errors greater than 100 cm). However, Superior and Ontario there are 510 possible configurations; and for Lake Michigan there are 254 possible configurations.

For all lakes, it is possible to pick three or four station networks with small errors in mean lake level (standard errors less than 0.5 cm and maximum errors less than 2.5 cm). These networks have stations near the midpoint of the lake or with each station near one end balanced by a station near the other end. However, for Lake Erie (Figure 17), it is also possible to pick a network that will produce mean lake levels with large errors (standard errors greater than 5 cm and maximum errors greater than 25 cm). These networks have most of their stations near one end of the lake. The first mode will be a combination of mean lake level and the lake surface tilt along the longer lake axis. For the other lakes, the worst case networks have much smaller errors due to the fact that the other lakes are much less responsive to wind stress than Lake Erie.

For each subnetwork size, the lowest standard error and the lowest maximum error are shown for each lake in Figure 22. It can be seen that only small reductions in mean lake level error are obtained for networks greater than seven stations. It should be noted that for each network size, there are numerous configurations that give errors only slightly larger than the lowest error subnetwork.

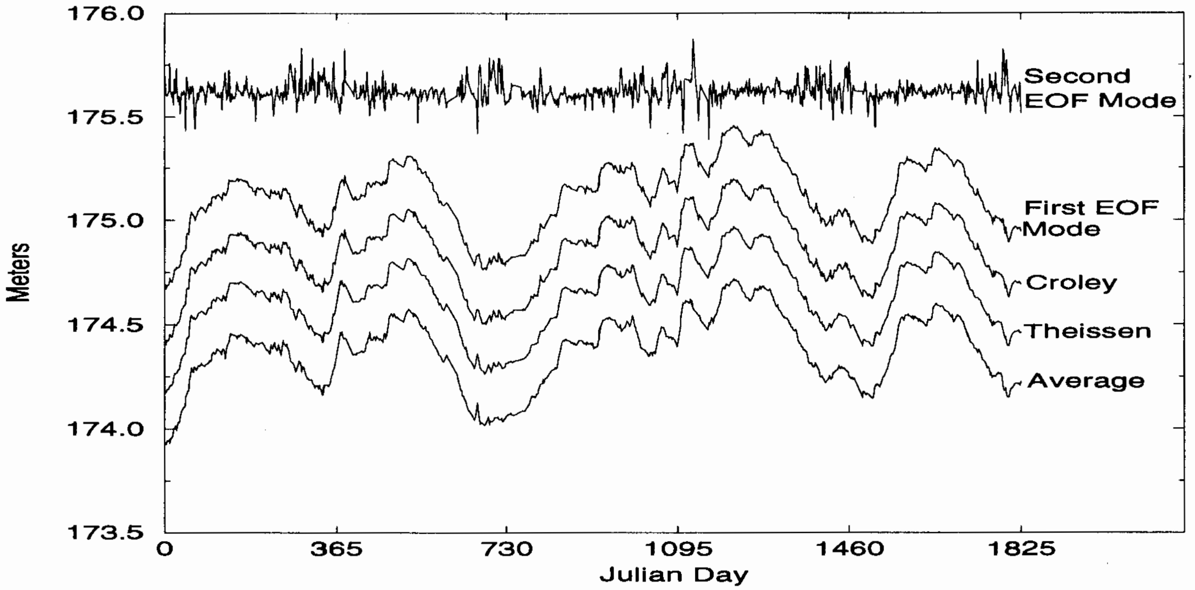


Figure 12. Mean lake levels (offset) for Lake Erie calculated for 2-day averaged water levels using uniform, Theissen, Croley, and EOF weights. Also shown is the second EOF mode.

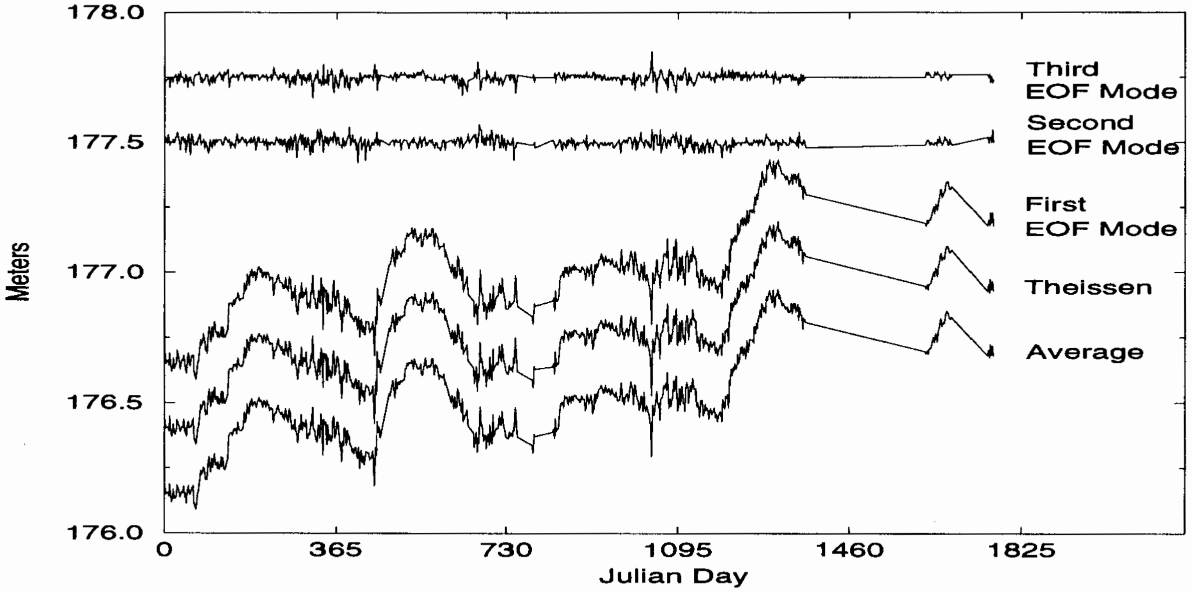


Figure 13. Mean lake levels (offset) for Lake Huron calculated for 2-day averaged water levels using uniform, Theissen, and EOF weights. Also shown are the second and third EOF modes.

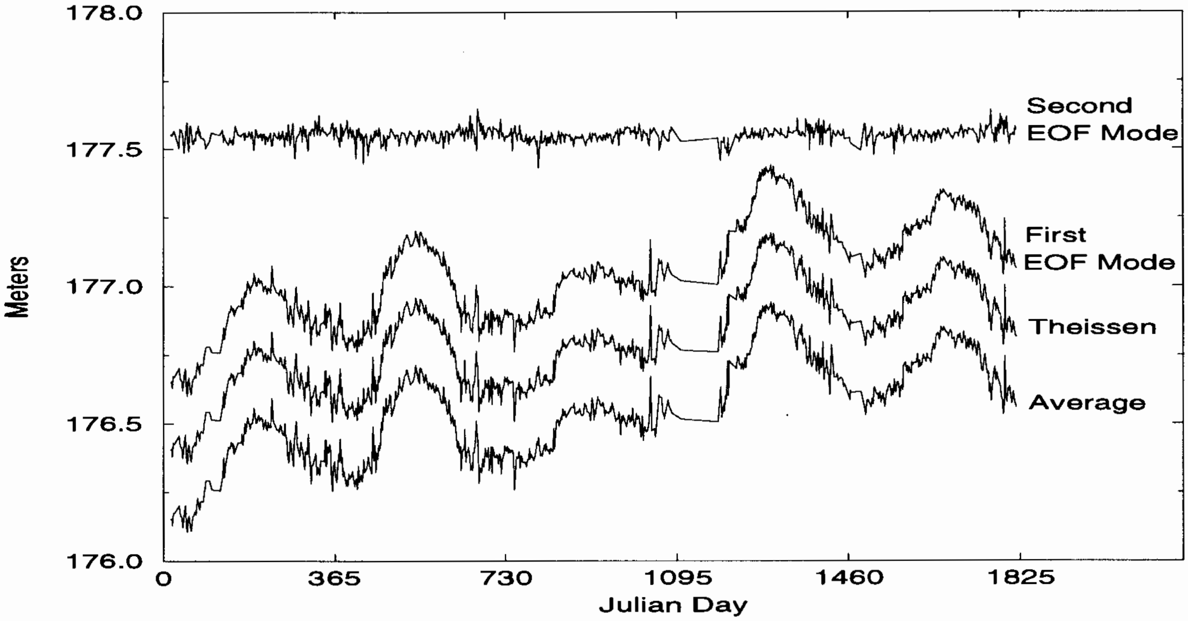


Figure 14. Mean lake levels (offset) for Lake Michigan calculated for 2-day averaged water levels using uniform, Theissen, and EOF weights. Also shown is the second EOF mode.

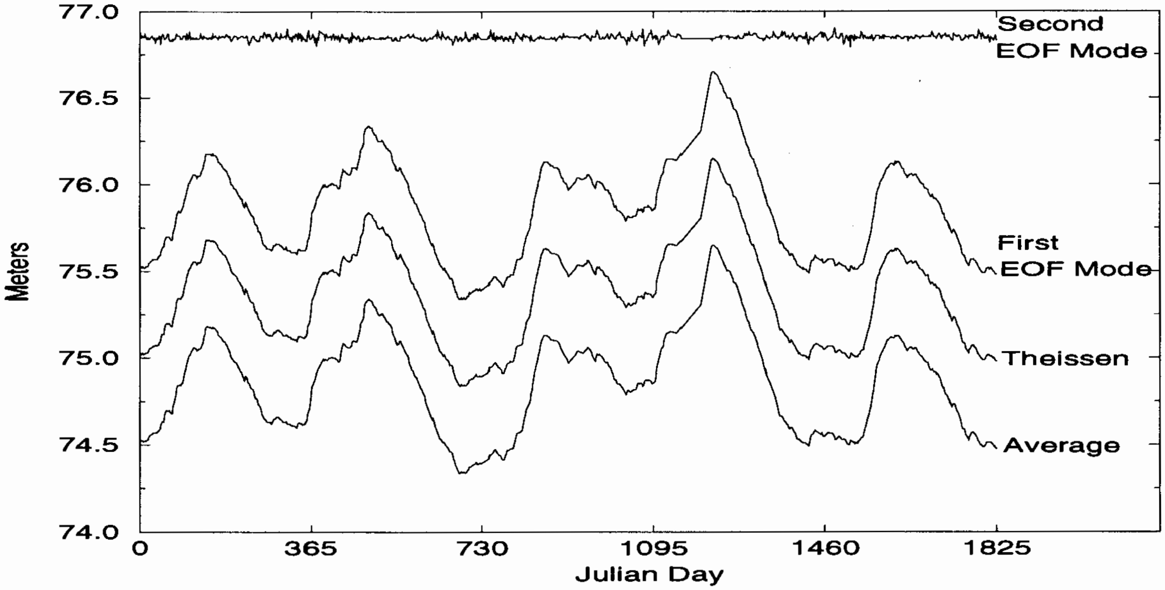


Figure 15. Mean lake levels (offset) for Lake Ontario calculated for 2-day averaged water levels using uniform, Theissen, and EOF weights. Also shown is the second EOF mode.

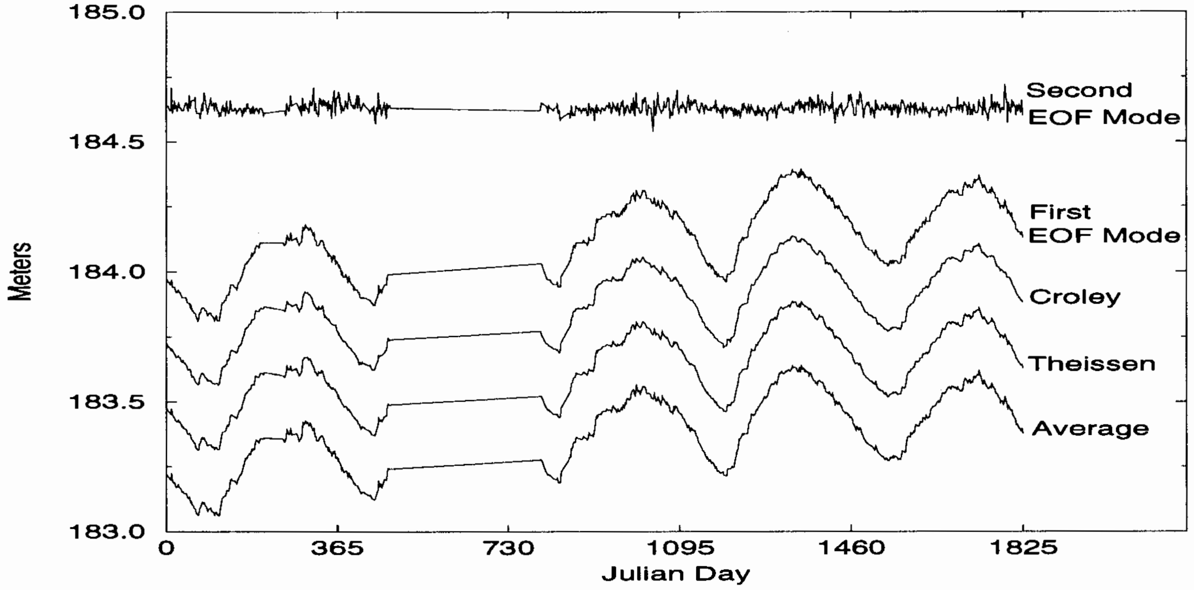
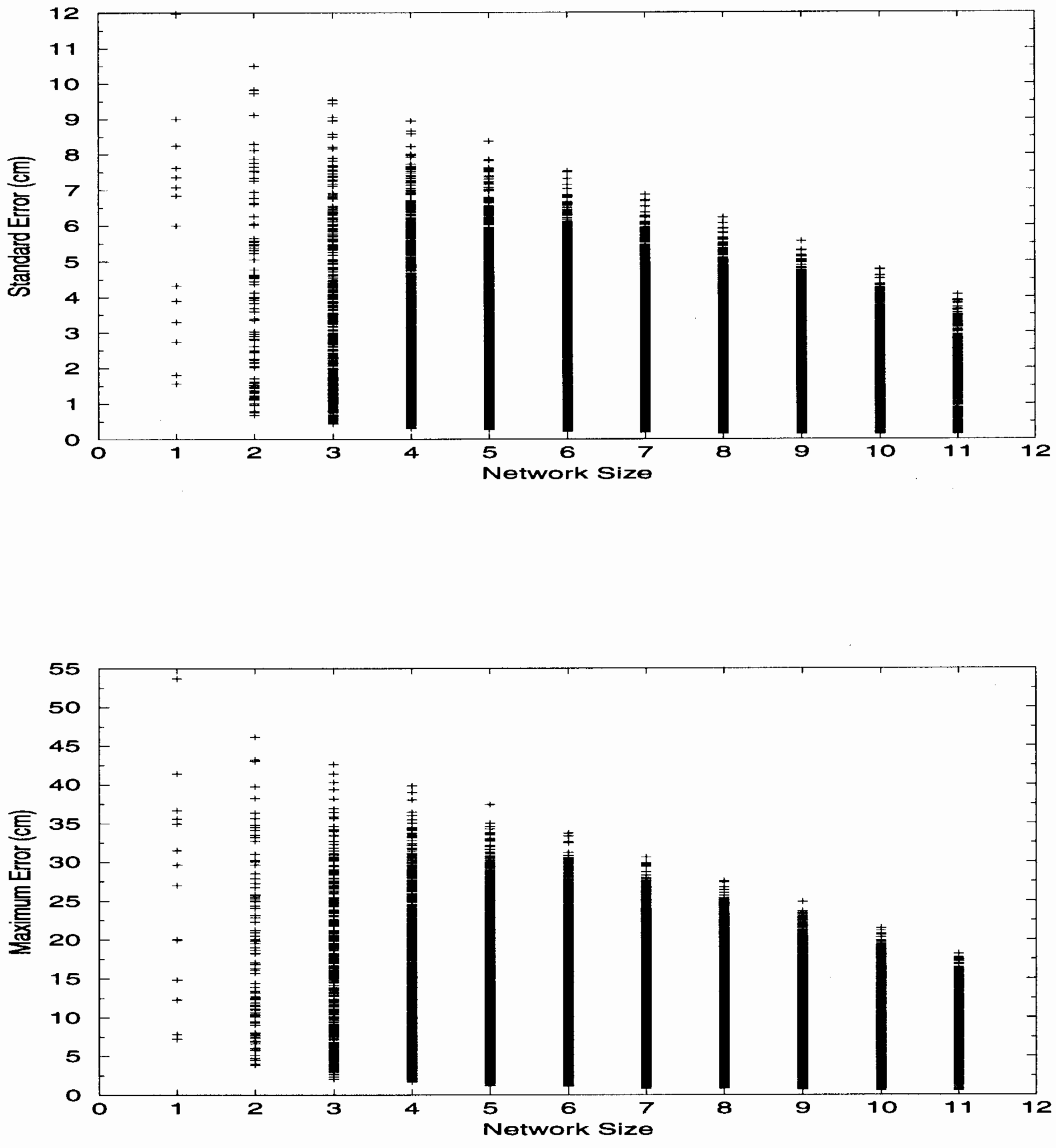


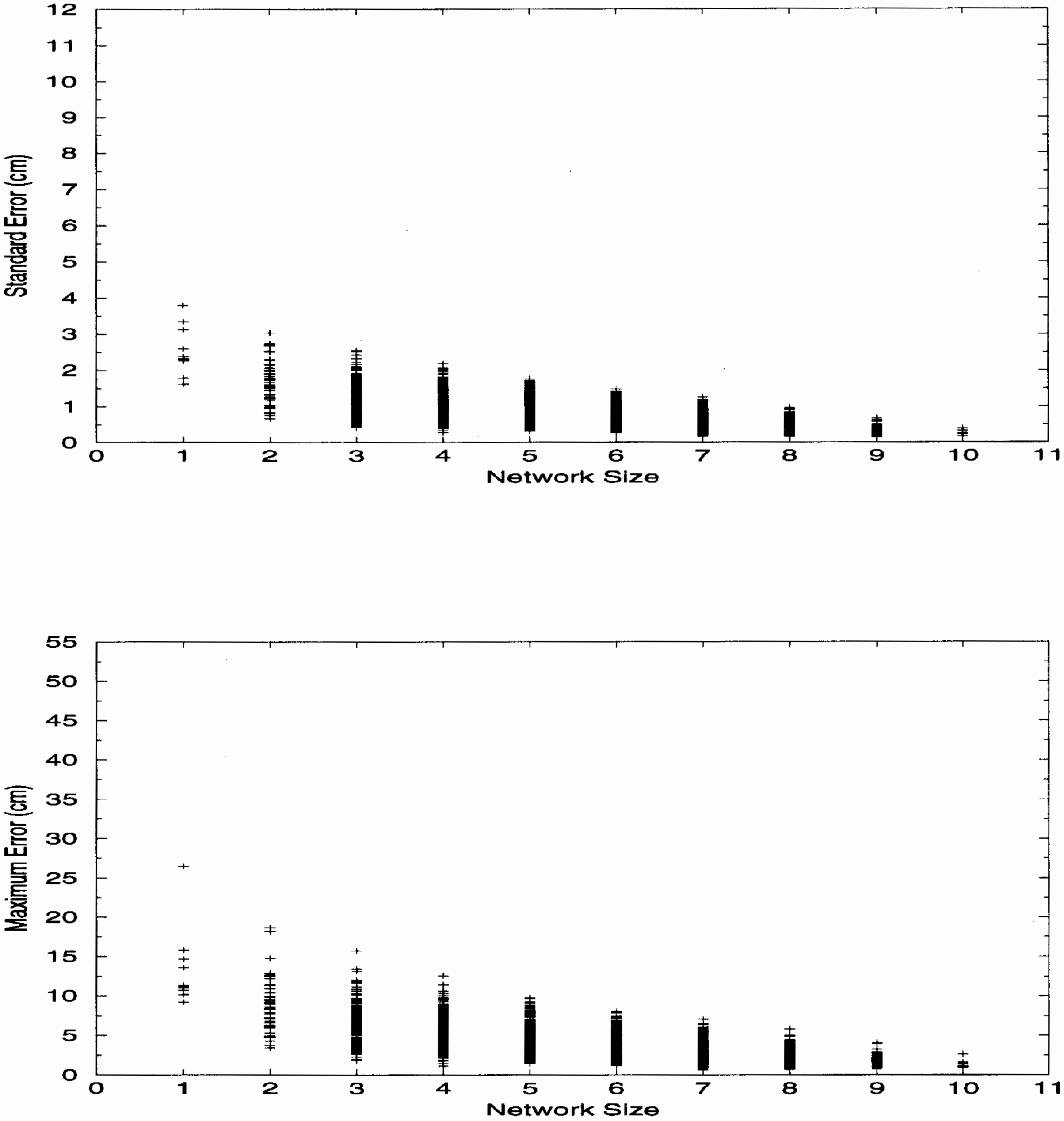
Figure 16. Mean lake levels (offset) for Lake Superior calculated for 2-day averaged water levels using uniform, Theissen, Croley, and EOF weights. Also shown is the second EOF mode.



a)

b)

Figure 17. a) Standard error and b) maximum error of subnetwork first EOF modes for Lake Erie 2-day averaged lake levels.



a)

b)

Figure 18. a) Standard error and b) maximum error of subnetwork first EOF modes for Lake Huron 2-day averaged lake levels.

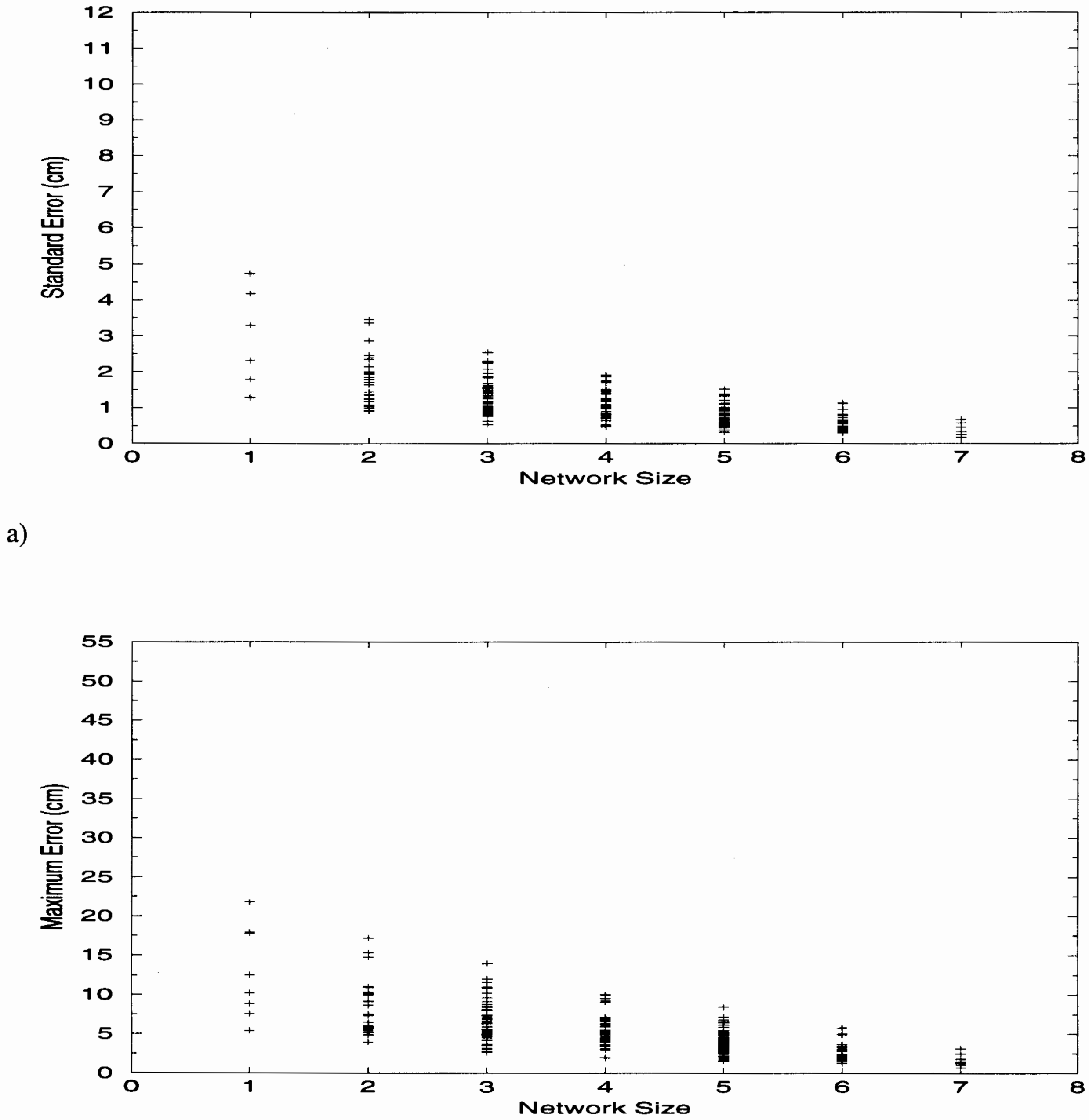


Figure 19. a) Standard error and b) maximum error of subnetwork first EOF modes for Lake Michigan 2-day averaged lake levels.

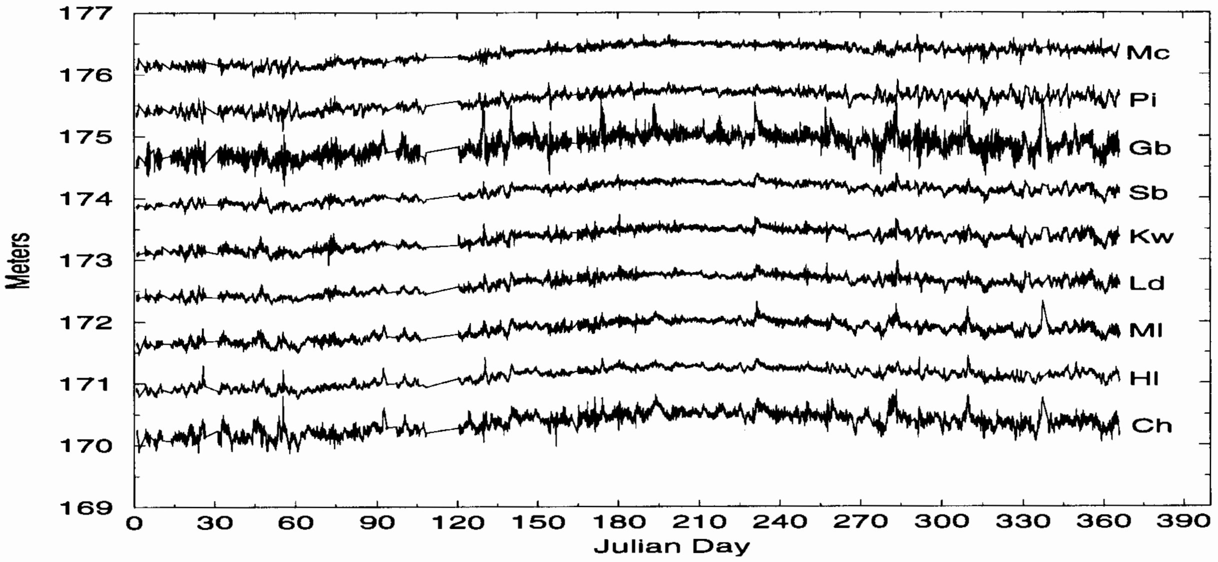


Figure 4. Lake Michigan hourly water levels for 1990 (offset for comparison). Mc - Mackinaw City, Pi - Port Inland, Gb - Green Bay, Sb - Sturgeon Bay, Kw - Kewaunee, Ld - Ludington, Ml - Milwaukee, Hl - Holland, Ch - Calumet Harbor.

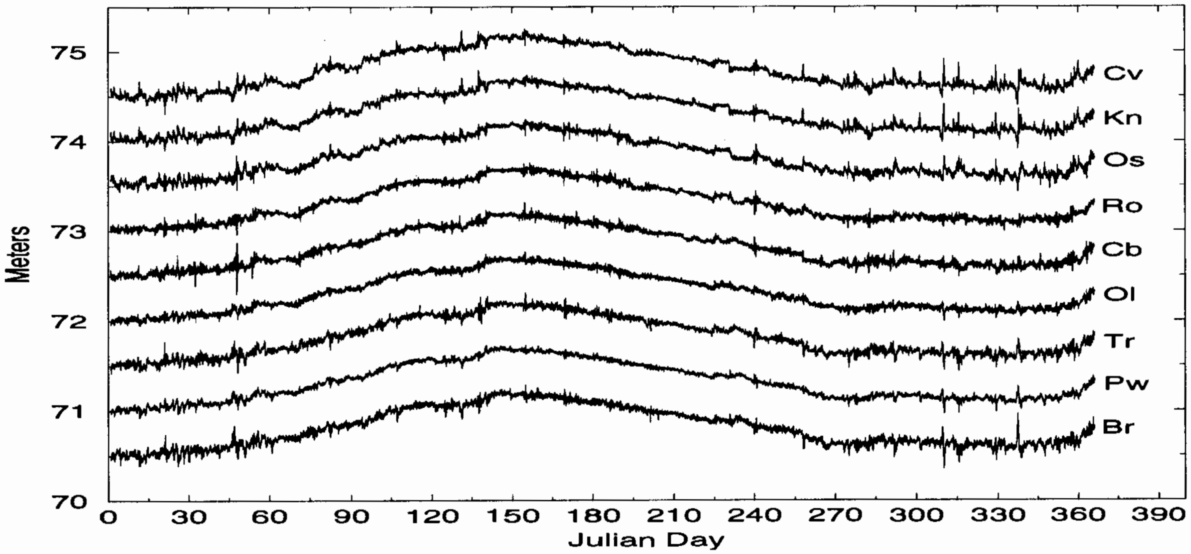


Figure 5. Lake Ontario hourly water levels for 1990 (offset for comparison). Cv - Cape Vincent, Kn - Kingston, Os - Oswego, Ro - Rochester, Cb - Cobourg, Ol - Olcott, Tr - Toronto, Pw - Port Weller, Br - Burlington.

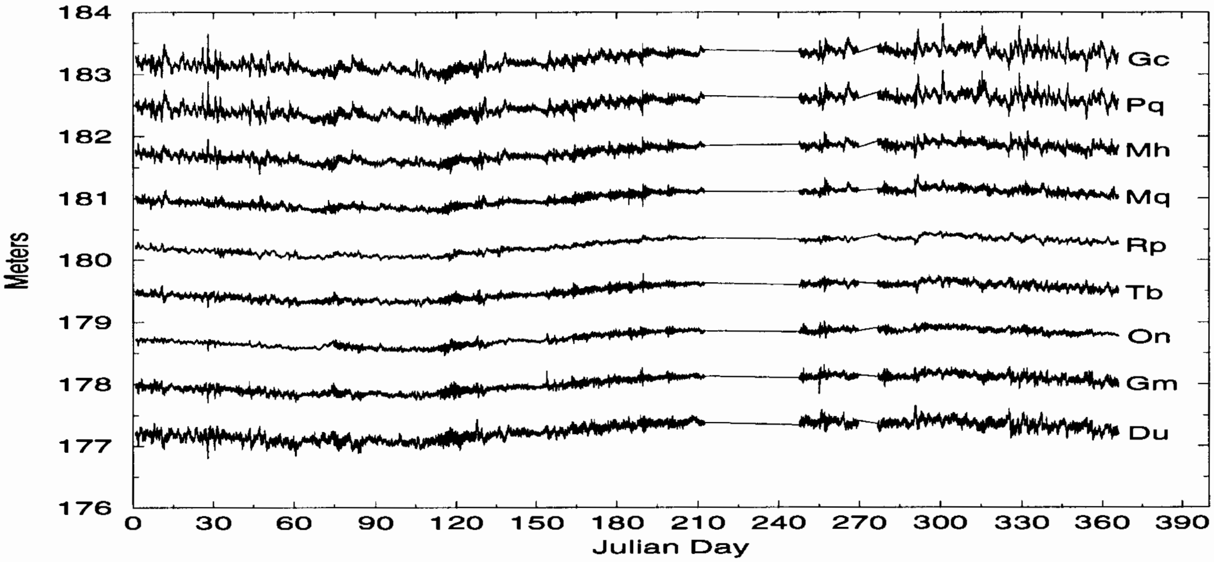


Figure 6. Lake Superior hourly water levels for 1990 (offset for comparison). Gc - Gros Cap, Pq - Port Iroquois, Mh - Michipicoten Harbour, Mq - Marquette, Rp - Rossport, Tb - Thunder Bay, On - Ontonagon, Gm - Grand Marais, Du - Duluth.

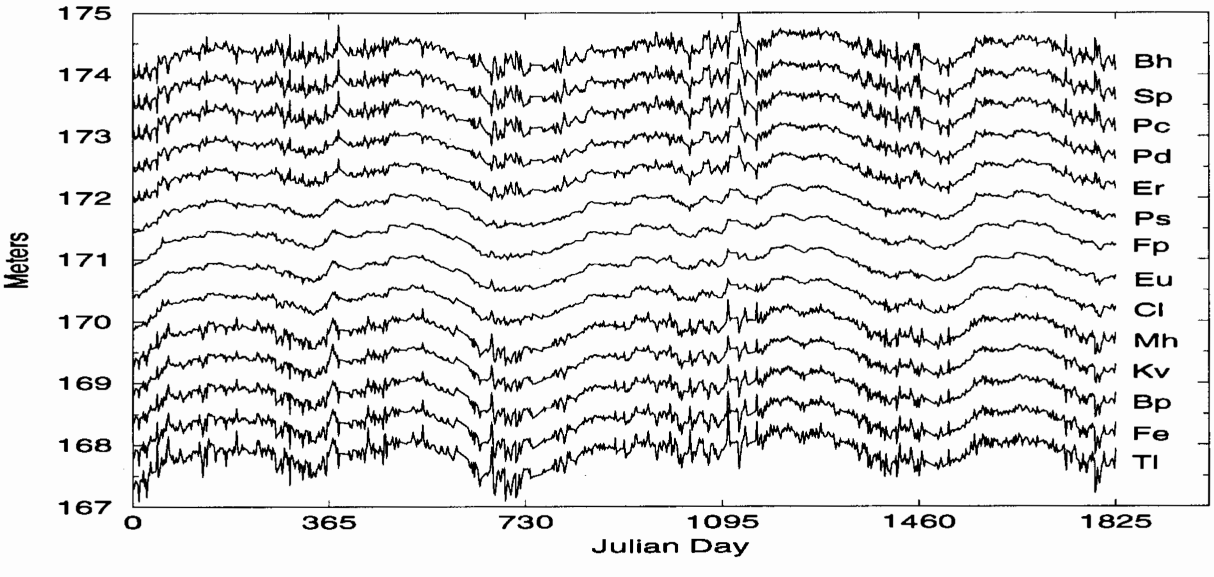


Figure 7. Lake Erie 2-day averaged water levels for 1990-1994 (offset for comparison). Bh - Buffalo Harbor, Sp - Sturgeon Point, Pc - Port Colborne, Pd - Port Dover, Er - Erie, Ps - Port Stanley, Fp - Fairport, Eu - Erieau, Cl - Cleveland, Mh - Marblehead, Kv - Kingsville, Bp - Bar Point, Fe - Fermi, Tl - Toledo.

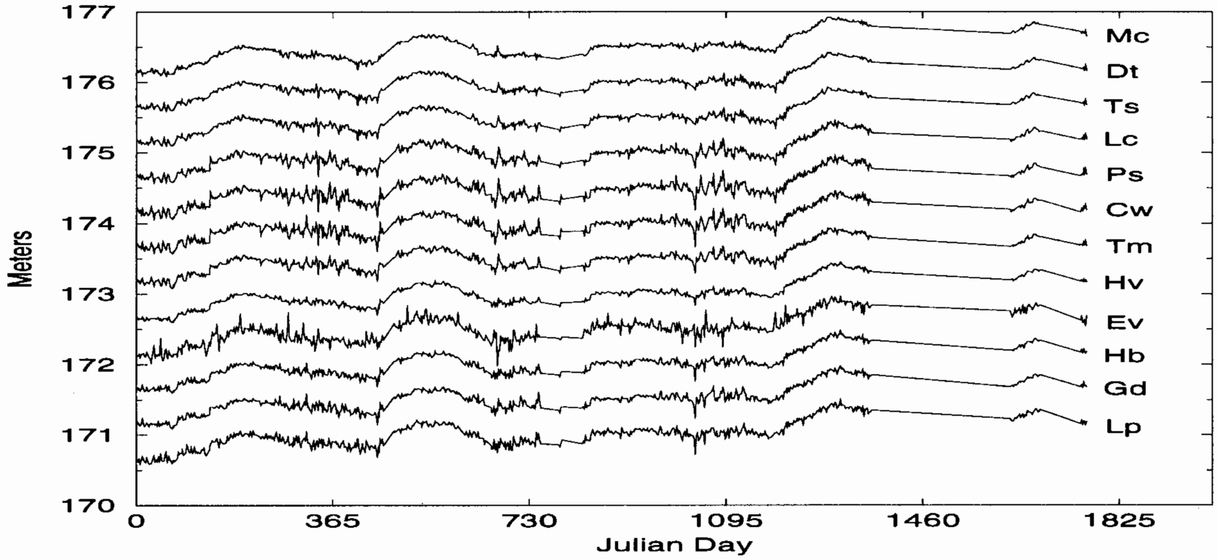


Figure 8. Lake Huron 2-day averaged water levels for 1990-1994 (offset). Mc - Mackinaw City, Dt - De Tour, Ts - Thessalon, Lc - Little Current, Ps - Parry Sound, Cw - Collingwood, Tm - Tobermory, Hv - Harrisville, Ev - Essexville, Hb - Harbor Beach, Gd - Goderich, Lp - Lakeport.

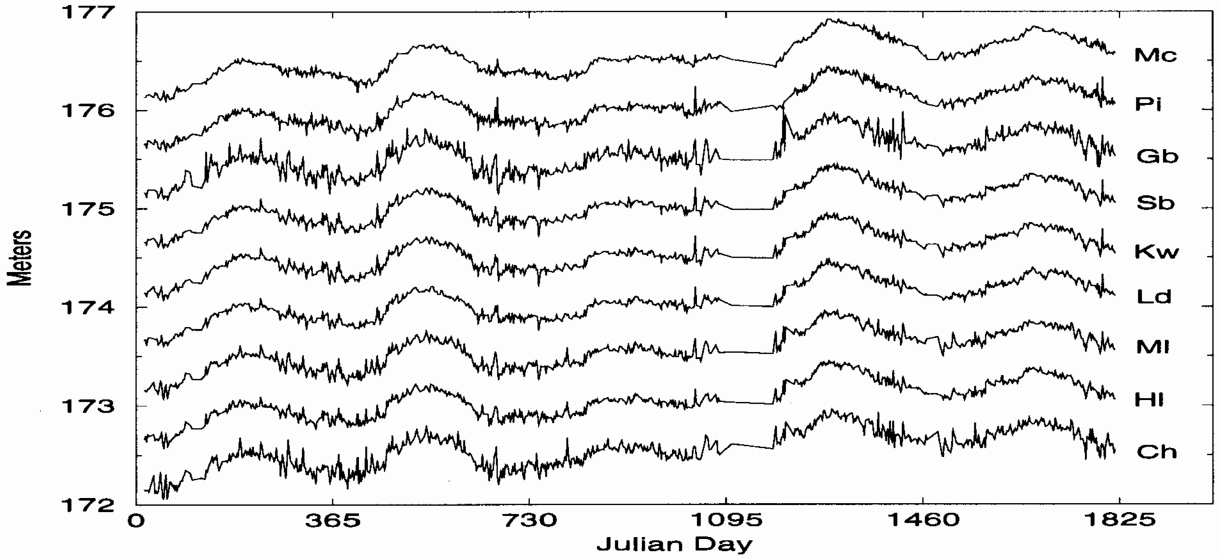


Figure 9. Lake Michigan 2-day averaged water levels for 1990-1994 (offset for comparison). Mc - Mackinaw City, Pi - Port Inland, Gb - Green Bay, Sb - Sturgeon Bay, Kw - Kewaunee, Ld - Ludington, Ml - Milwaukee, Hl - Holland, Ch - Calumet Harbor.

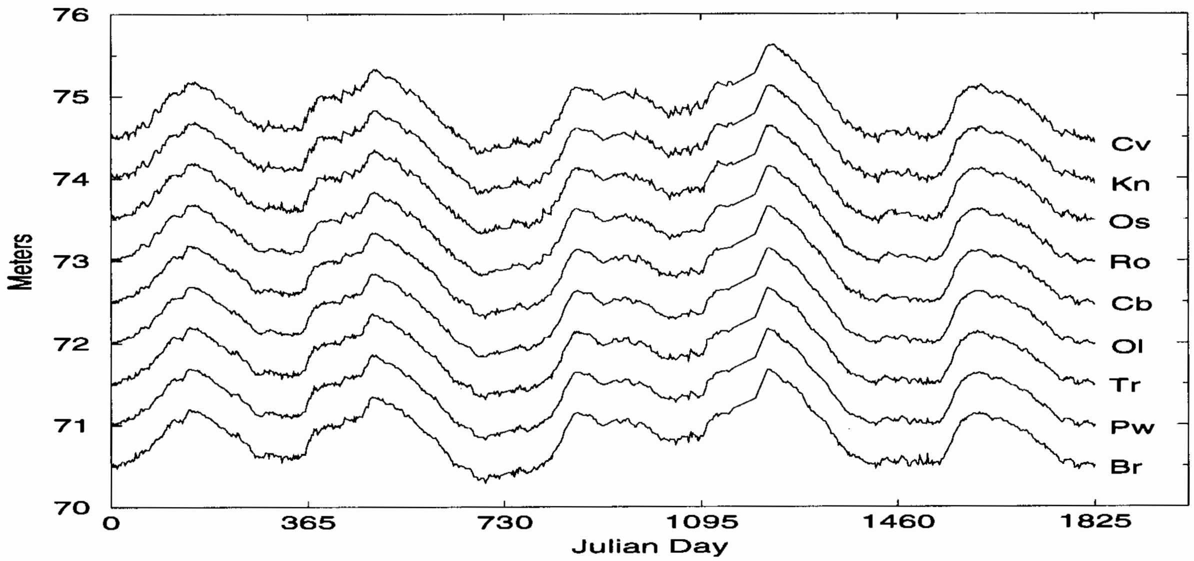


Figure 10. Lake Ontario 2-day averaged water levels for 1990-1994 (offset for comparison). Cv - Cape Vincent, Kn - Kingston, Os - Oswego, Ro - Rochester, Cb - Cobourg, Ol - Olcott, Tr - Toronto, Pw - Port Weller, Br - Burlington.

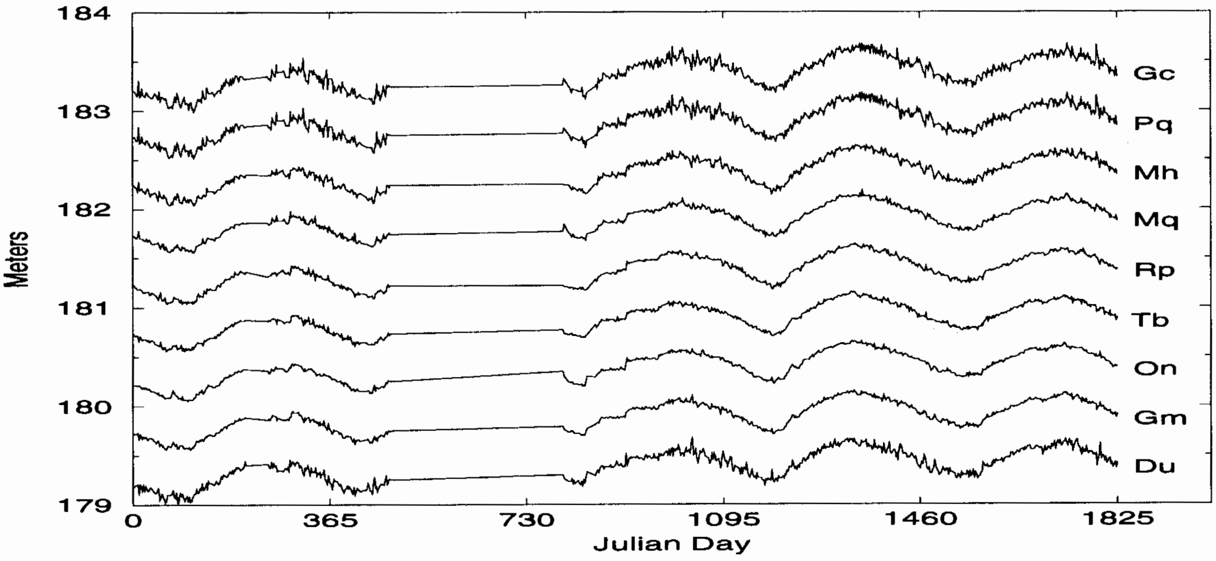


Figure 11. Lake Superior 2-day averaged water levels for 1990-1994 (offset for comparison). Gc - Gros Cap, Pq - Port Iroquois, Mh - Michipicoten Harbour, Mq - Marquette, Rp - Rossport, Tb - Thunder Bay, On - Ontonagon, Gm - Grand Marais, Du - Duluth.

4. ANALYSIS OF 2-DAY AVERAGED DATA

4.1. MEAN LAKE LEVEL COMPARISONS

The four different methods of approximating the mean lake level are shown in Figures 12 to 16 (offset for comparison). The Croley weights were only available for Lakes Erie and Superior. All four methods produce similar mean lake level signals, primarily composed of an annual cycle with a range of 40 to 60 cm and an interannual variability. Lakes Ontario and Erie generally reach their highest levels in the spring, Lakes Huron and Michigan usually peak in the summer, and Lake Superior peaks in the fall. There also appears to be a consistent rise in lake levels for Lakes Superior, Michigan, and Huron during this 5-year period. It is easily seen that Lakes Huron and Michigan mean lake levels have greater high frequency variability than the other lakes, an observation that will be discussed later.

Table 1. Mean Lake Level Difference Statistics (cm)
relative to the Theissen Method
1990-1994 2-day averaged data

	Method	Standard error	Maximum error
Lake Erie	Average	0.62	3.10
	Croley	0.65	3.70
	EOF	0.83	3.16
Lake Huron	Average	0.48	2.30
	EOF	0.81	3.09
Lake Michigan	Average	0.31	2.20
	EOF	0.39	2.27
Lake Ontario	Average	0.29	1.20
	EOF	0.29	1.24
Lake Superior	Average	0.32	1.50
	Croley	0.37	2.20
	EOF	0.32	1.19

Since a time series of the true mean lake level is not available, we can only compare the methods with each other. The Theissen method generally gives a slightly smoother curve since the quieter stations near the middle of the lake are more heavily weighted than stations near the ends of the lake which are more affected by wind set-up. If the Theissen method is assumed to produce the closest approximation to real mean lake levels (as the Corps of Engineers does), difference statistics relative to the Theissen method (standard error and maximum error) may be calculated for the other methods (Table 1). All errors are small (less than 1 cm standard error), with the largest errors for Lakes Erie and Huron. The Theissen weights for Lake Huron (Quinn, 1975b) were obtained using only two Canadian stations on the east side of the lake (Collingwood and Goderich) and therefore may not give the best approximation of the mean lake level. Although the EOF method does not always have the lowest errors, it provides more insight into the main physical processes operating in the lakes.

4.2. MODAL ENERGY DISTRIBUTION

When an EOF analysis was carried out for the full network for each lake, the first mode contained the greatest part of the energy ranging from 90.2% for Lake Erie to 99.7% for Lake Ontario (Table 2). The second mode, with energy percentages ranging from 9.3% for Lake Erie to 0.2% for Lake Ontario, is the lake surface tilt along the longer axis of each lake (i.e. the eigenvectors show that the second mode at one end of the lake is 180 degrees out of phase with the other end). The second EOF mode is shown along with the first EOF mode in Figures 12 to 16. The relative energy of the second mode to the first mode can be observed. The third EOF mode is also shown for Lake Huron since it is comparable in amplitude to the second.

Table 2. EOF Mode Energy Percentage (2-Day Averaged Data)

Mode	1	2	3	≥4
Superior	97.4	2.0	0.4	0.2
Michigan	97.5	1.8	0.3	0.4
Huron	98.0	1.0	0.8	0.2
Erie	90.2	9.3	0.3	0.2
Ontario	99.7	0.2	0.0	0.1

Lake Erie is the shallowest of the Great Lakes and therefore is most responsive to wind stress. The lake surface tilt along the shorter axis of the lake is the third mode for Lakes Superior, Huron, and Ontario and the fourth mode for Lakes Michigan and Erie. The third mode for Lakes Michigan and Erie is a mode with both ends of the lake out of phase with the middle of the lakes. Only for Lake Huron, due to its irregular shape, is the third mode comparable in energy to the second mode.

4.3. NETWORK SIZE

Since the first EOF mode obtained from the complete network is a good measure of the mean lake level, we now statistically compare the first EOF mode obtained from smaller networks to the first EOF mode obtained from the complete network. This is done by examining the standard error and the maximum error for every possible subnetwork (Figures 17 to 21). For the Lake Erie there are 16,382 possible configurations; for Lake Huron there are 2046 possible configurations; for Lakes more spatially balanced network configurations have small errors (standard errors less than 1 cm and maximum errors less than 10 cm). For the other lakes, the worst case networks have errors much smaller than the worst case networks of Lake Erie.

The lowest standard errors and the lowest maximum errors for each subnetwork size are shown in Figure 33. It can be seen that only small error reductions are obtained for networks greater than eight stations. It should be kept in mind that for each network size, a large number of other configurations will give only marginally greater errors.

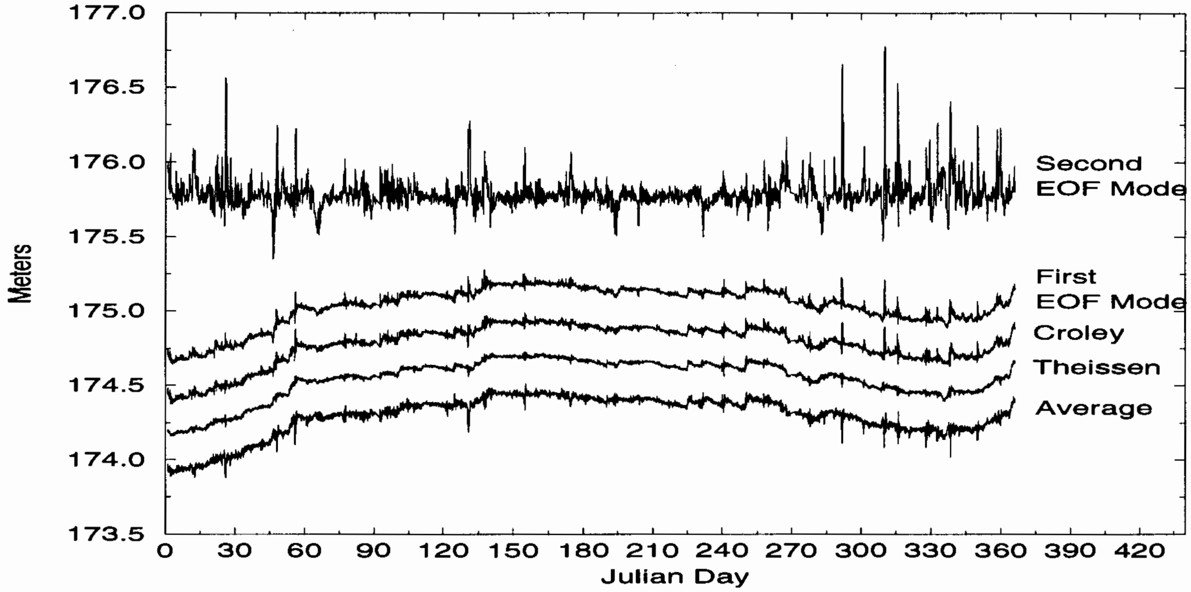


Figure 23. Mean lake levels (offset) for Lake Erie calculated for hourly water levels for 1990 using uniform, Theissen, Croley, and EOF weights. Also shown is the second EOF mode.

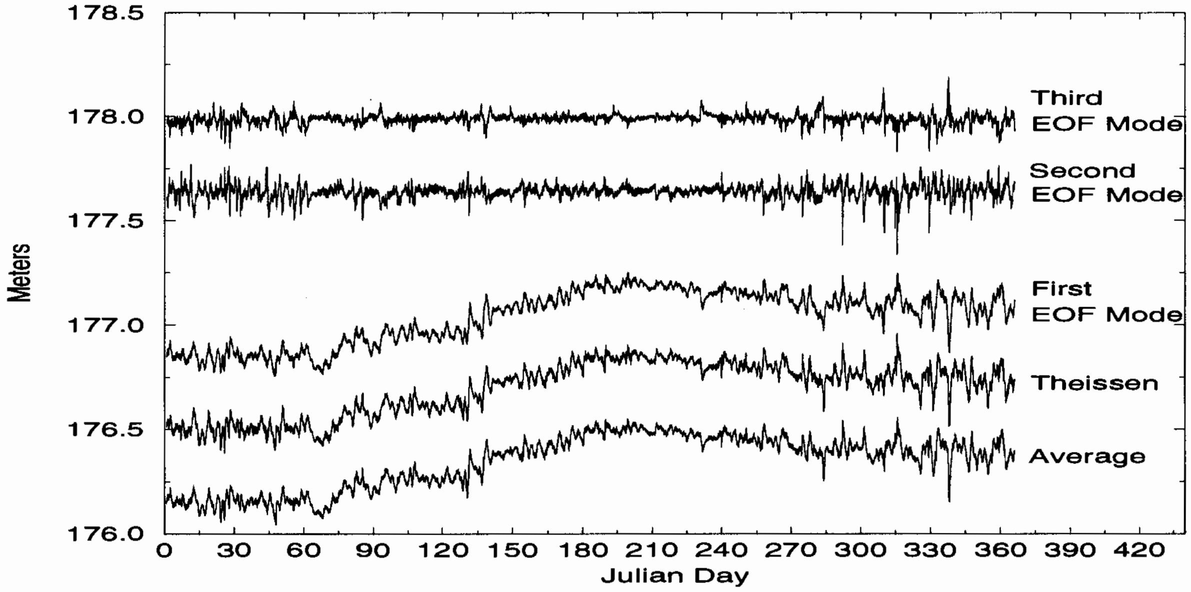


Figure 24. Mean lake levels (offset) for Lake Huron calculated for hourly water levels for 1990 using uniform, Theissen, and EOF weights. Also shown are the second and third EOF mode.

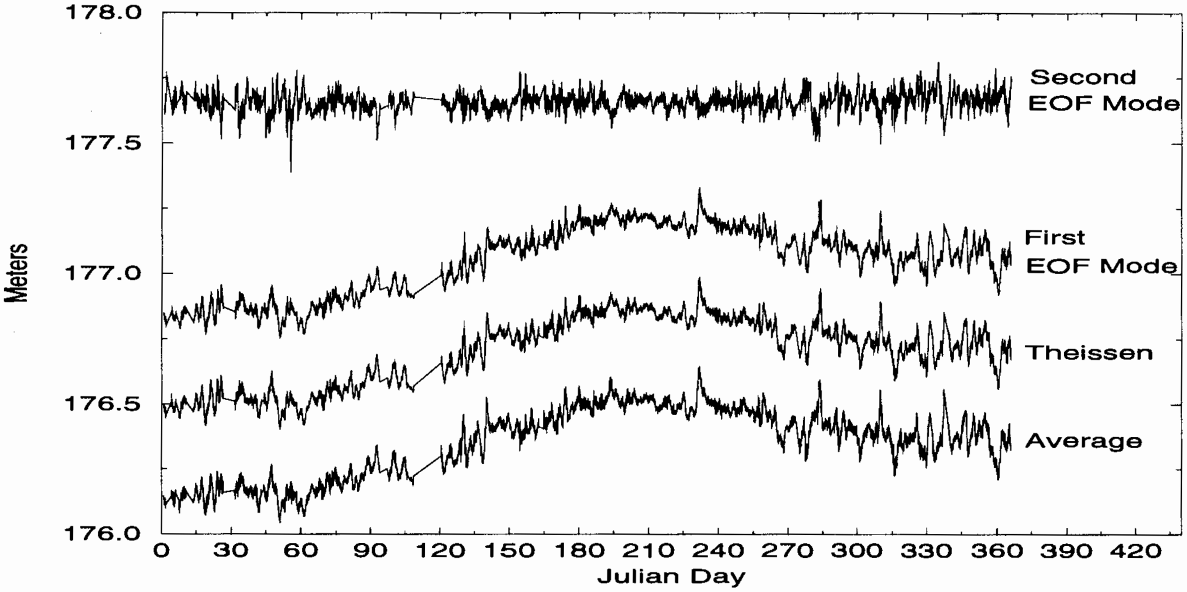


Figure 25. Mean lake levels (offset) for Lake Michigan calculated for hourly water levels for 1990 using uniform, Theissen, and EOF weights. Also shown is the second EOF mode.

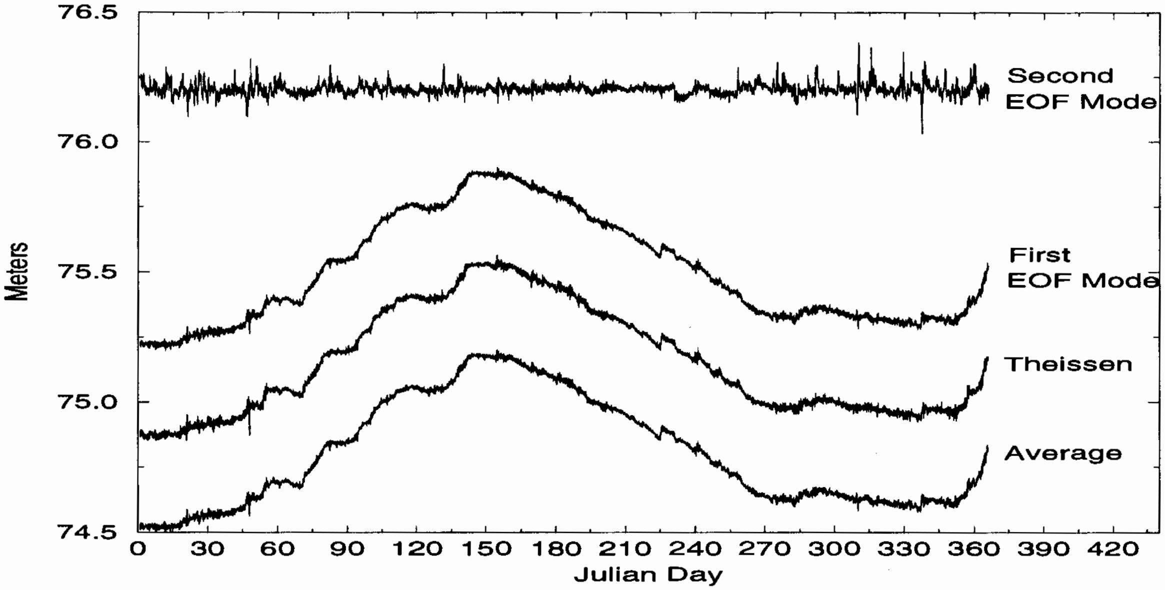


Figure 26. Mean lake levels (offset) for Lake Ontario calculated for hourly water levels for 1990 using uniform, Theissen, and EOF weights. Also shown is the second EOF mode.

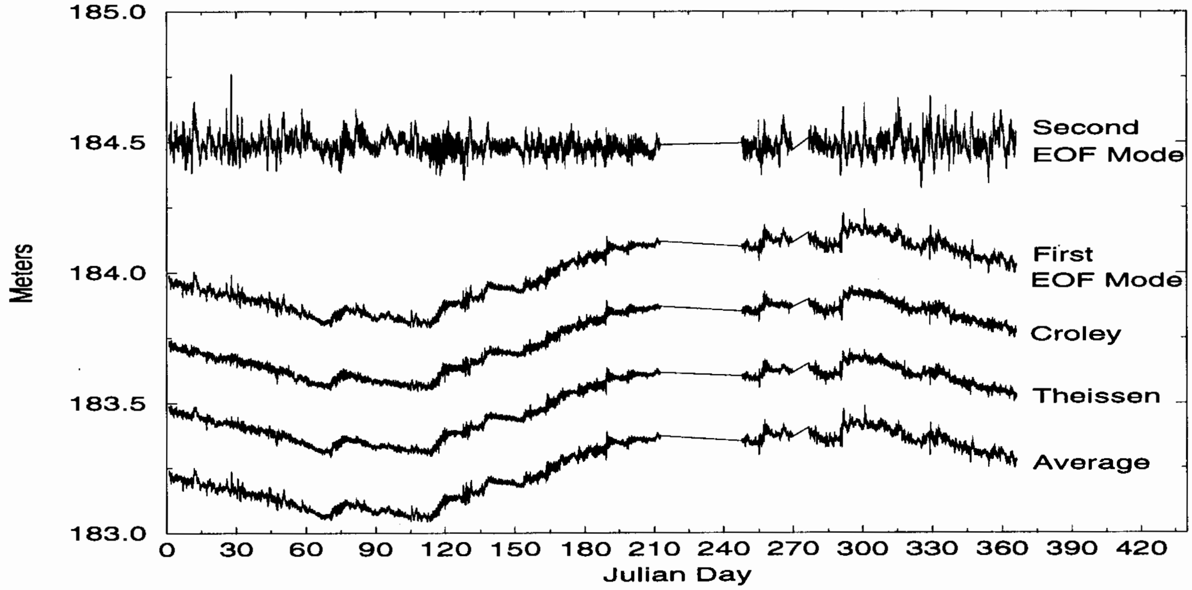
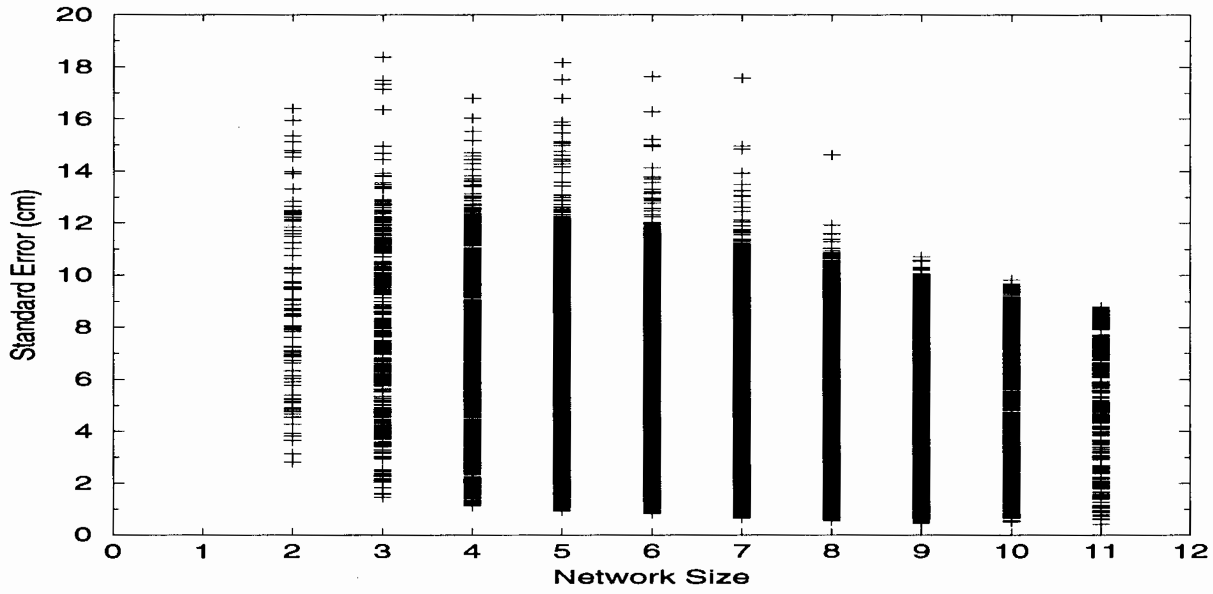
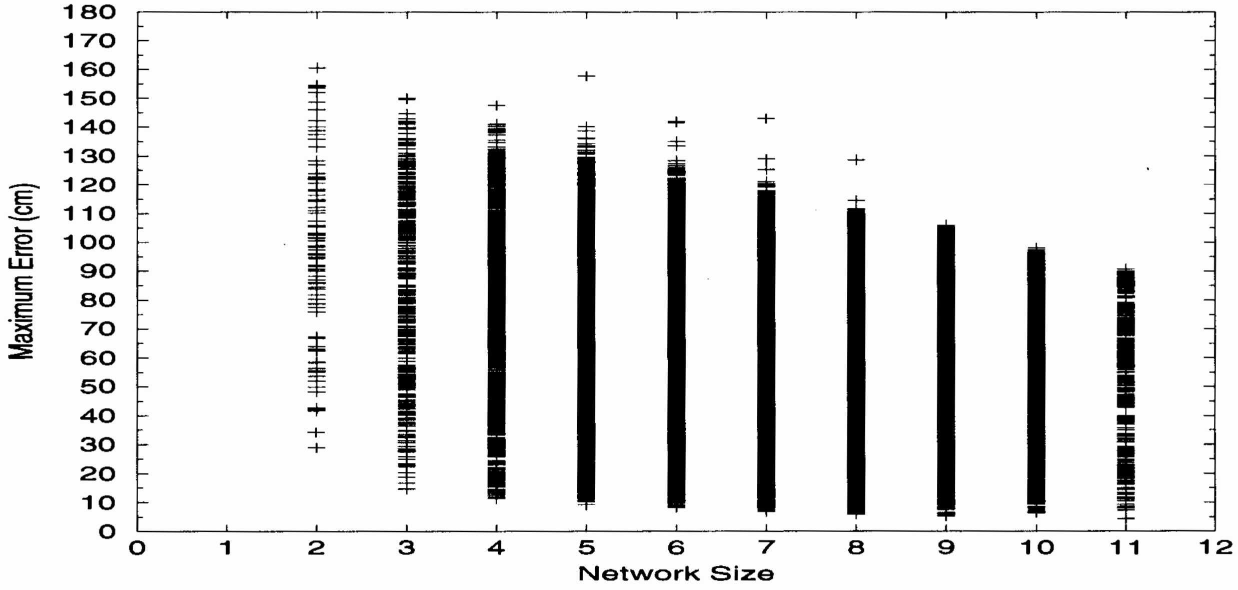


Figure 27. Mean lake levels (offset) for Lake Superior calculated for hourly water levels for 1990 using uniform, Theissen, Croley, and EOF weights. Also shown is the second EOF mode.



a)



b)

Figure 28. a) Standard error and b) maximum error of subnetwork first and second EOF modes for Lake Erie hourly water levels.

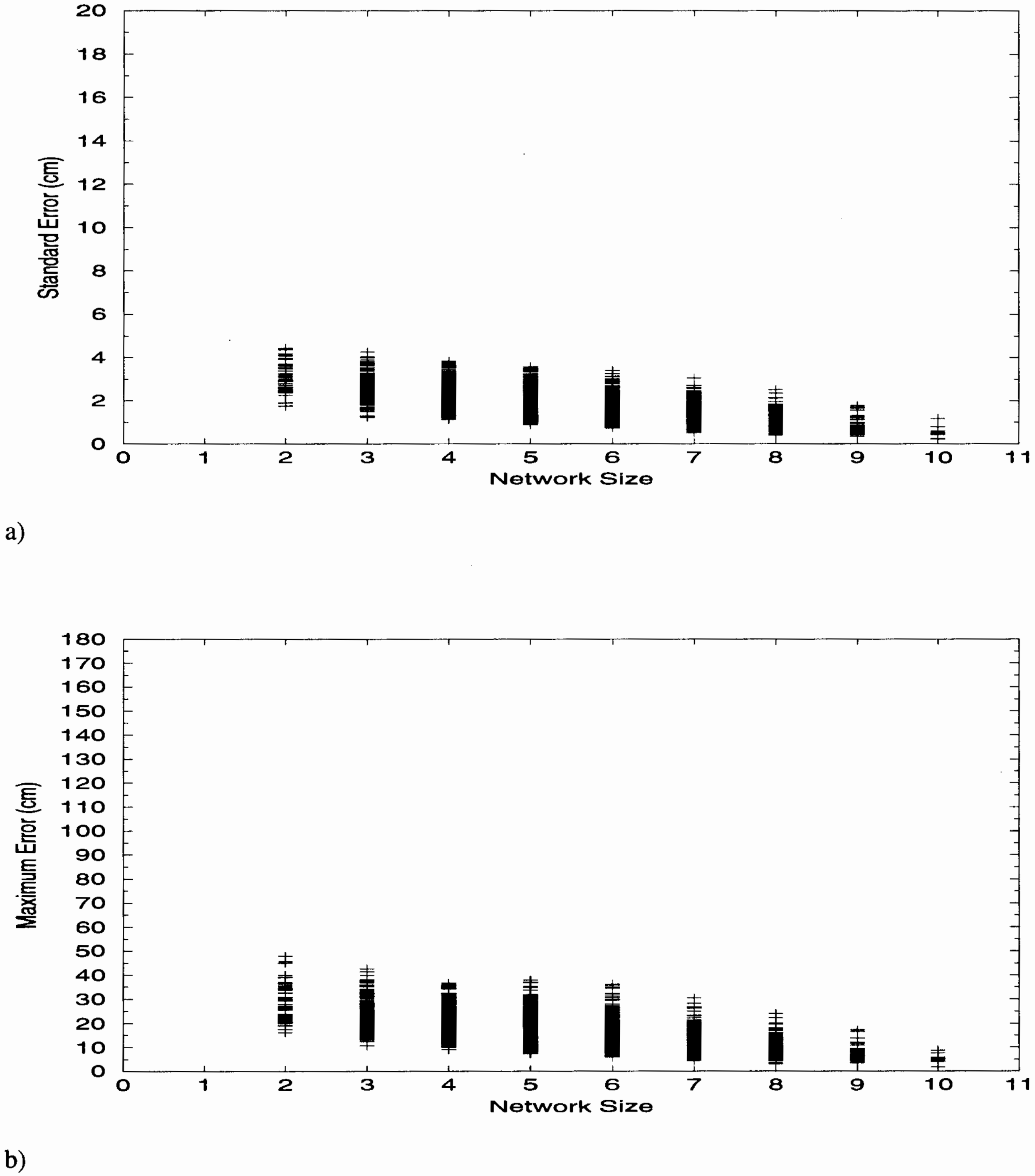
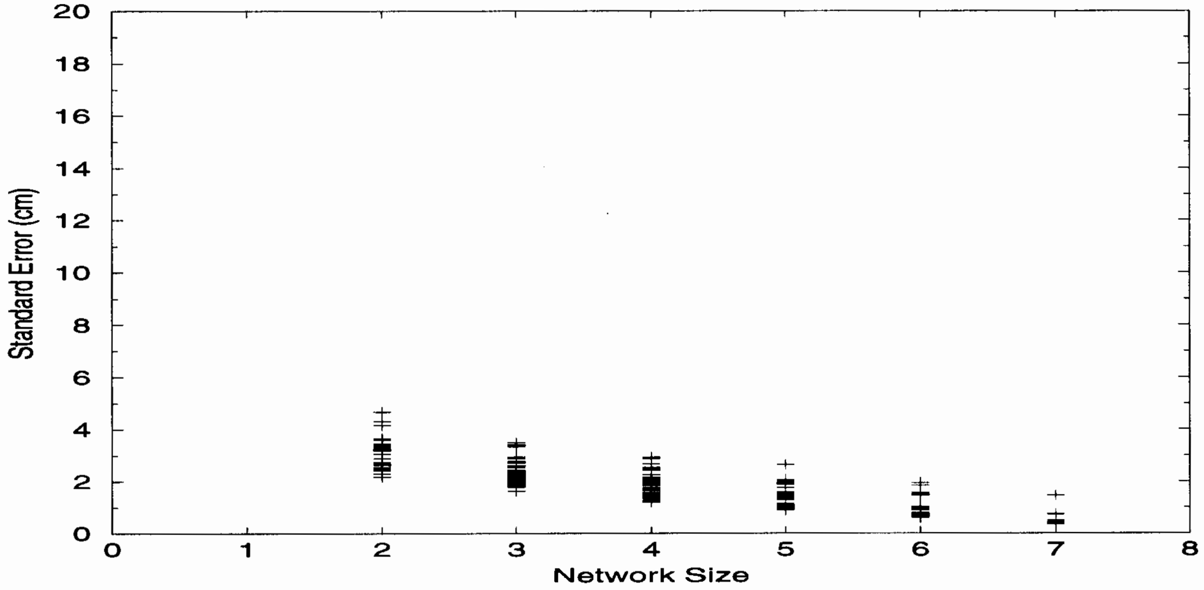
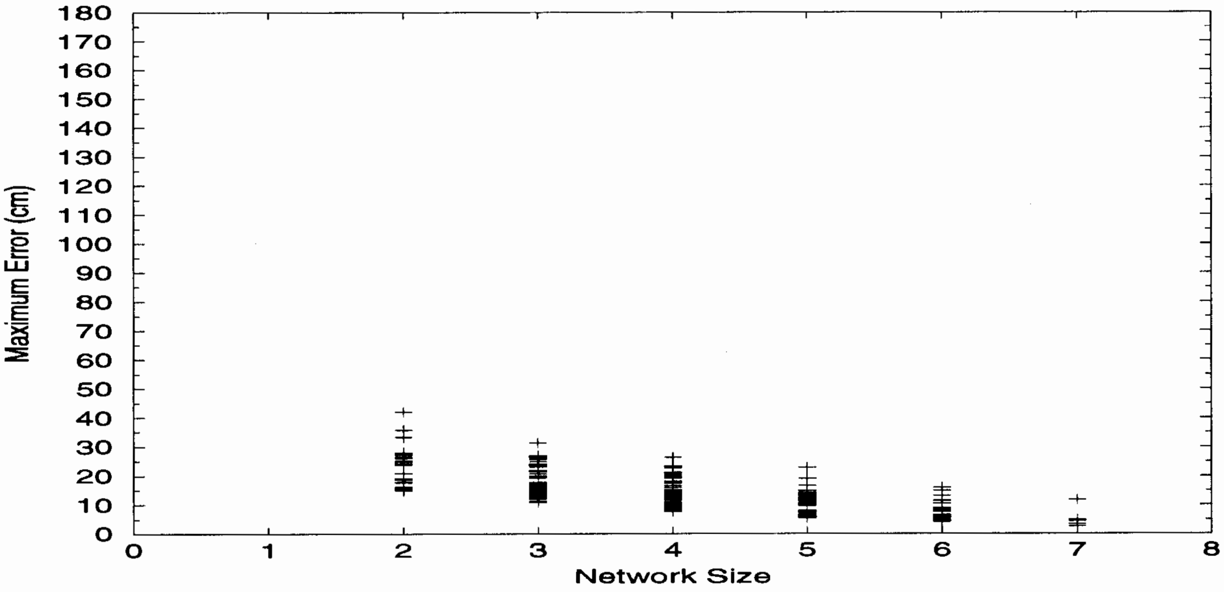


Figure 29. a) Standard error and b) maximum error of subnetwork first and second EOF modes for Lake Huron hourly water levels.

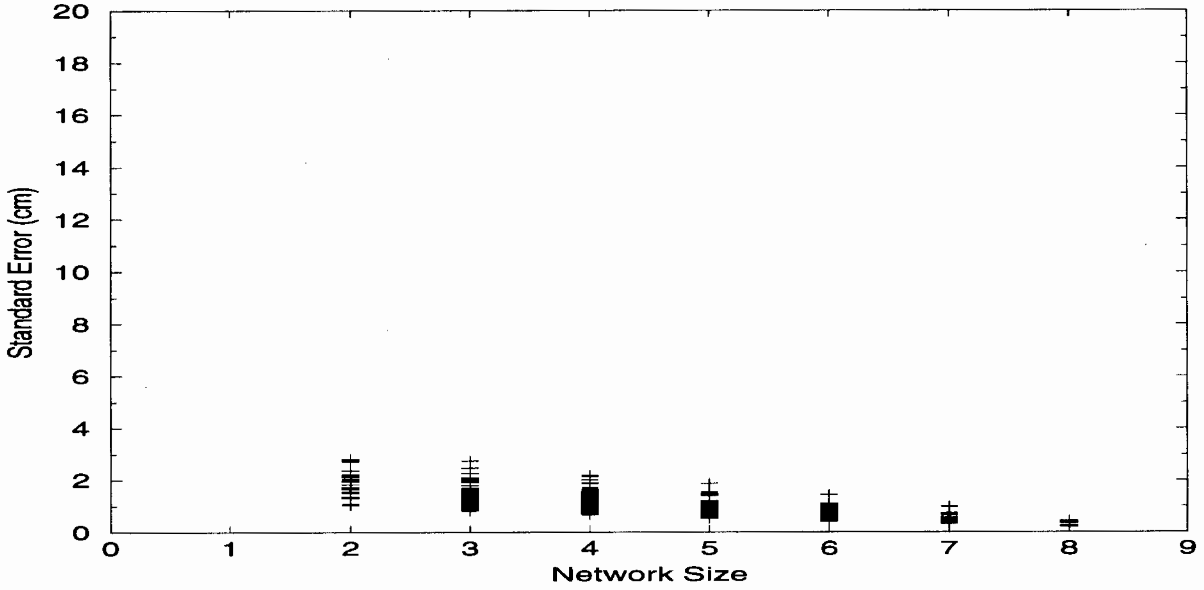


a)

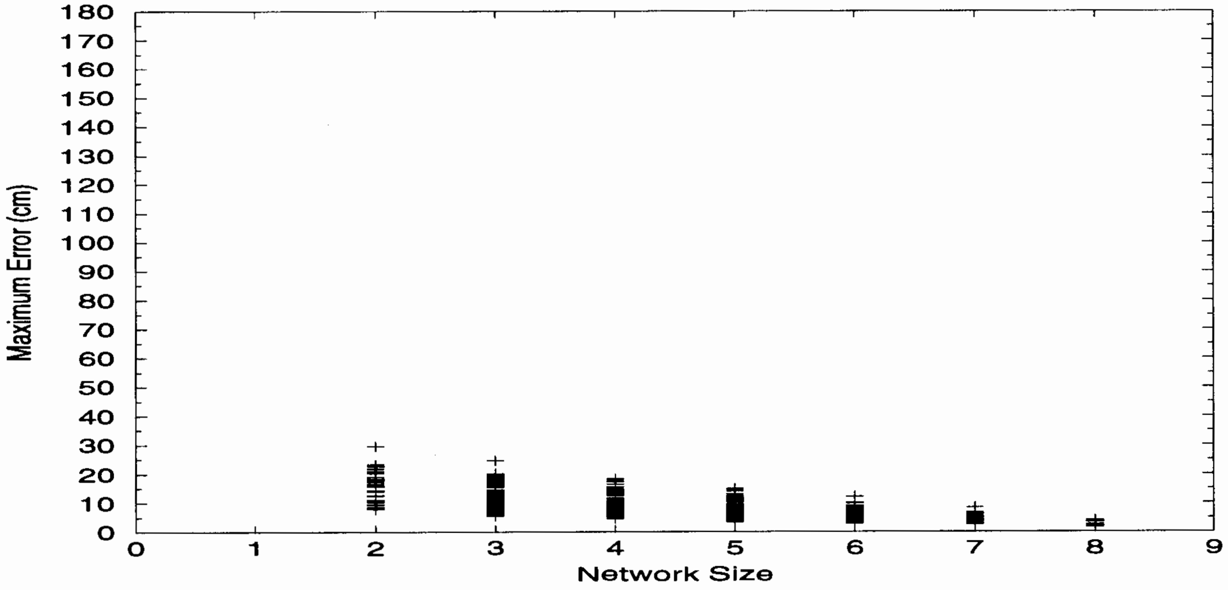


b)

Figure 30. a) Standard error and b) maximum error of subnetwork first and second EOF modes for Lake Michigan hourly water levels.

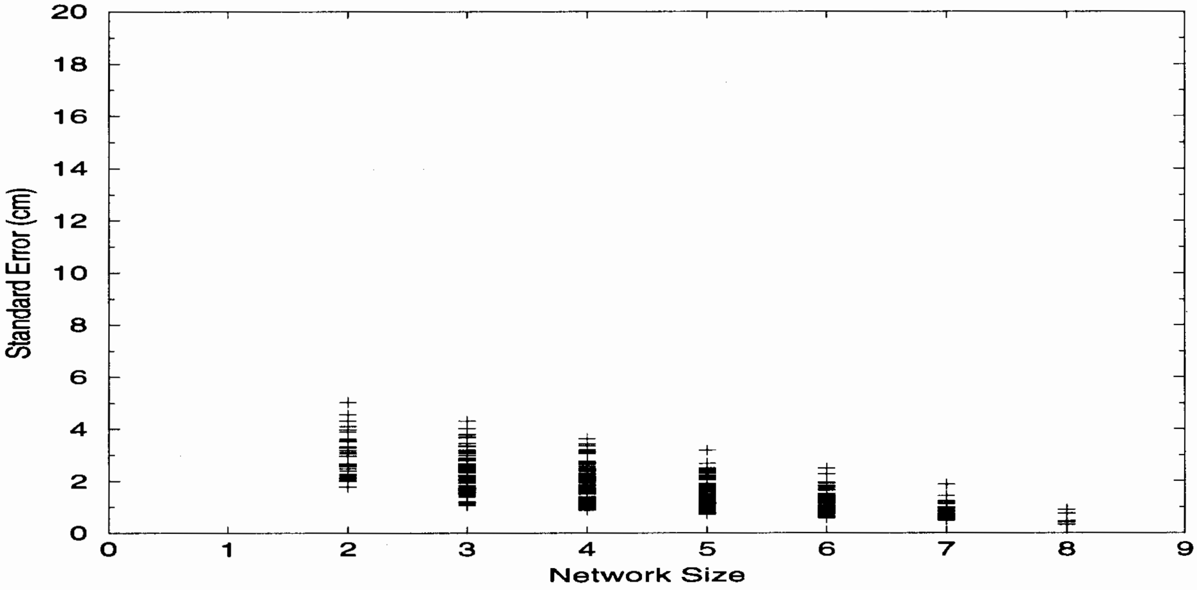


a)

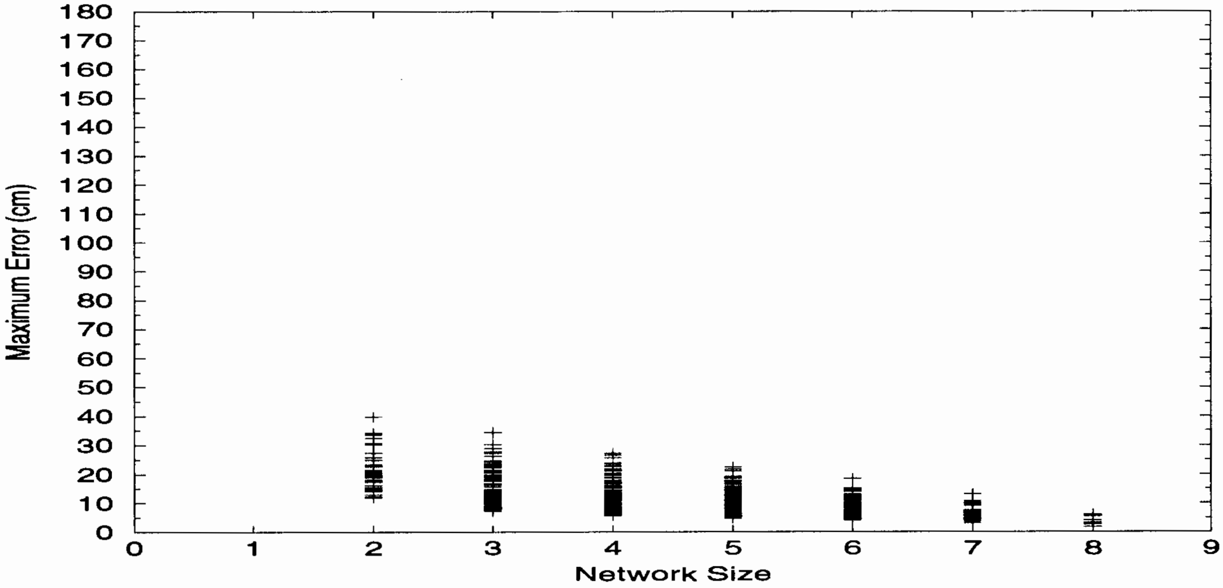


b)

Figure 31. a) Standard error and b) maximum error of subnetwork first and second EOF modes for Lake Ontario hourly water levels.



a)



b)

Figure 32. a) Standard error and b) maximum error of subnetwork first and second EOF modes for Lake Superior hourly water levels.

7. CONCLUSION

The goal of this investigation was to quantify the reduction in mean lake level accuracy that would occur if the number of gauges presently in place on the Great Lakes were to be reduced. Five years of water level data (1990-1994) from 52 stations on the Great Lakes were used in the analyses of network configurations. First, the calculation of the mean lake level using 2-day averaged data for each lake was examined. This statistic is required by the Corps of Engineers for the regulation of lake levels. An empirical orthogonal function analysis for each lake showed that 90.2% to 99.7% of the energy was in the first mode corresponding to the mean lake level. The first EOF mode obtained with smaller network configurations was then compared with the first EOF mode from the complete network. The difference statistics indicate the accuracy of each subnetwork. The lowest standard errors and lowest maximum errors decrease with increasing subnetwork size up to a seven station network with a standard error below 0.2 cm and a maximum error below 1 cm.

Next, the EOF analysis was repeated for hourly data for 1990. Without the time averaging of the data, the first mode (corresponding to mean lake level) had a smaller percentage of the total energy (61.7% to 98.0%). The second mode (corresponding to lake surface tilt along the longer axis of each lake) had a larger percentage of the energy (1.1% to 34.0%). The EOF analysis was then carried out for every possible subnetwork, and the difference statistics for both the first and the second mode were examined. The lowest standard errors and lowest maximum errors decreased with increasing network size up to an eight station network with a standard error below 0.7 cm and a maximum error below 7 cm.

If a station is dropped from the network, it is possible to obtain an approximation of the signal at that location using other stations still in the network. This can be done by using the other stations to calculate the time series for the first and second EOF modes and then scaling them for the station not in the reduced network. Alternatively, the high correlation coefficients between stations that are close together suggest simply substituting the time series of the nearest station. As an example, the four, six, and eight station subnetworks with the lowest standard error were used to approximate stations not in the subnetworks. In some cases, reconstructing the signal gave a better approximation while in other cases, substituting the nearest station gave a better approximation.

ACKNOWLEDGEMENTS

The author wishes to thank a number of people who assisted in the course of this investigation. Philip Morris, Brooks Widder, and Jeff Oyler of the NOS/Ocean and Lake Levels Division provided the U.S. water level data and obtained the Canadian water level data. Charles Sun wrote the EOF analysis program. The report was significantly improved by the reviews provided Kurt Hess, Charles Sun, Edward Shih, Tom Eisler, and Dave Pendleton. The initial plan for this investigation was developed in cooperation with Bruce Parker, Eugene Russin, and Philip Morris.

REFERENCES

Clites, A. H., Editor, 1992. Improved communication of Great Lakes water level information, NOAA Tech. Memo. ERL GLERL-77, 80 pp.

Coordinating Committee, 1995.* Establishment of International Great Lakes Datum (1985), Coordinating Committee of Great Lakes Basic Hydraulic and Hydrologic Data, 60 pp.

Croley II, T. E., 1986. Minimizing long-term wind set-up errors in estimated mean Erie and Superior lake levels. Great Lakes Environ. Res. Lab., Ann Arbor, MI, GLERL Contrib. 535, 45 pp.

Croley II, T. E., 1987. Wind set-up error in mean lake levels. J. Hydrol., 92: 223-243.

Croley II, T. E., and H. C. Hartmann, 1985. Resolving Theissen polygons, J. Hydrol. 76, 363-379.

David, M. H., E. F. Joeres, E. D. Loucks, K. W. Potter, and S. S. Rosenthal, 1988. Effects of diversions on the North American Great Lakes, Water Resour. Bull., vol. 24, no. 1, 141-148.

Grima, A. P., and C. Wilson-Hodges, 1977. Regulation of Great Lakes water levels: The public speaks out, J. Great Lakes Research, 3(3-4):240-257.

Hamblin, P. F., 1987. Meteorological forcing and water level fluctuations on Lake Erie, J. Great Lakes Res., 13 (4), 436-453.

Hartmann, H. C., 1988. Potential variation of Great Lakes water levels: A hydrologic response analysis, NOAA Tech. Memo. ERL GLERL-68, 30 pp.

Kundu, P. K., J. S. Allen, and R. L. Smith, 1975. Modal decomposition of the velocity field near the Oregon coast, J. Phys. Oceanogr., 5, 683-704.

National Ocean Service, 1994. Great Lakes Water Levels 1994, U. S. Dept. Of Commerce, NOAA, NOS, Silver Spring, MD, 206 pp.

Preisendorfer, R. W., 1988. Principal Component Analysis in Meteorology and Oceanography, Developments in Atmospheric Science 17, Elsevier, 425 pp.

Quinn, F. H., 1975a. Lake Michigan beginning-of-month water levels and monthly rates of change of storage. U. S. Gov. Print. Off., Washington, D.C., NOAA Tech. Rep. ERL 326-GLERL 2, 32 pp.

Quinn, F. H., 1975b. Lake Huron beginning-of-month water levels and monthly rates of change of storage. U. S. Gov. Print. Off., Washington, D.C., NOAA Tech. Rep. ERL 348-GLERL 4, 27 pp.

Quinn, F. H., and J. A. Derecki, 1976a. Lake Erie beginning-of-month water levels and monthly rates of change of storage. U. S. Gov. Print. Off., Washington, D.C., NOAA Tech. Rep. ERL 364-GLERL 9, 34 pp.

Quinn, F. H., and J. A. Derecki, 1976b. Lake Ontario beginning-of-month water levels and monthly rates of change of storage. U. S. Gov. Print. Off., Washington, D.C., NOAA Tech. Rep. ERL 365-GLERL 10, 27 pp.

Quinn, F. H., and M. J. Todd, 1974. Lake Superior beginning-of-month water levels and monthly rates of storage changes. U. S. Gov. Print. Off., Washington, D.C., NOAA Tech. Memo. NOS LSC R4, 38 pp.

Tovell, W. M. 1979. The Great Lakes: geology at work, J. Soil Water Conserv., 34(2), 65-67.

APPENDIX

Table A. Best Substitute Station Error Statistics (cm)

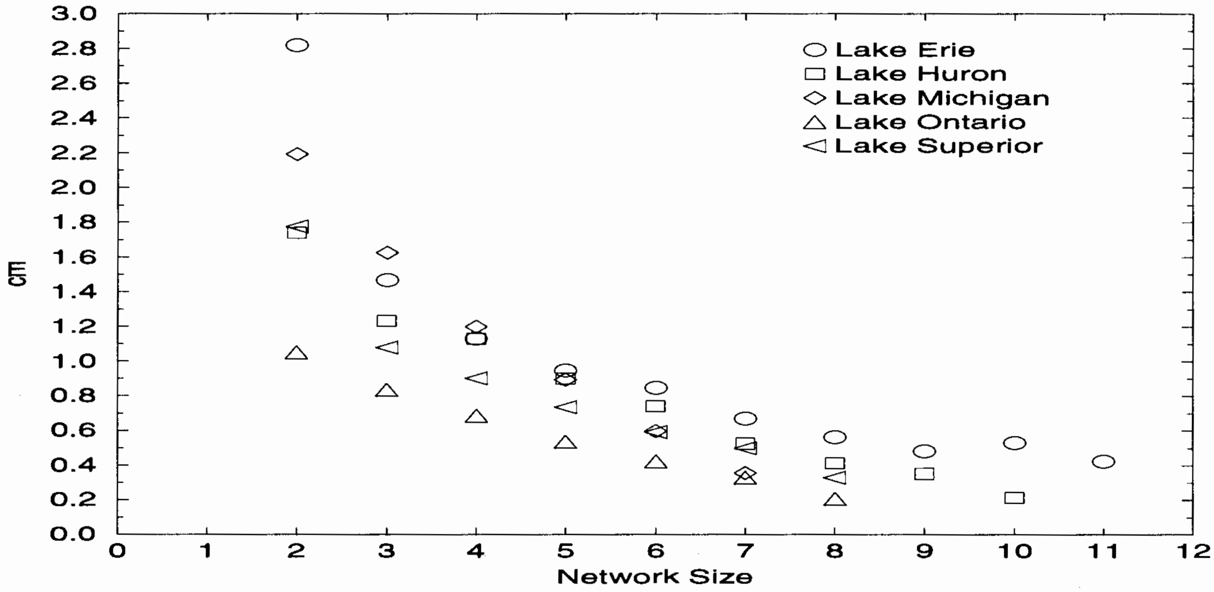
	Station	Substitute	Mean Error	Standard Error	Maximum Error
Lake Erie	Buffalo	Sturgeon Point	0.8	3.4	39
	Sturgeon Point	Port Colborne	-1.1	2.6	30
	Port Colborne	Sturgeon Point	1.1	2.6	30
	Port Dover	Erie	-1.3	4.1	35
	Erie	Port Dover	1.3	4.1	35
	Port Stanley	Erieau	1.7	6.3	58
	Fairport	Cleveland	2.0	4.2	46
	Erieau	Fairport	-2.0	5.2	69
	Cleveland	Fairport	-2.0	4.2	46
	Marblehead	Kingsville	-1.1	5.3	57
	Kingsville	Bar Point	0.5	4.9	44
	Bar Point	Fermi	0.6	3.5	45
	Fermi	Bar Point	-0.6	3.5	45
	Toledo	Fermi	-0.4	6.9	77
Lake Huron	Mackinaw City	De Tour	0.2	3.5	30
	De Tour	Thessalon	-0.3	2.1	17
	Thessalon	De Tour	0.3	2.1	17
	Little Current	Tobermory	-2.5	4.9	43
	Parry Sound	Collingwood	-0.7	3.5	30
	Collingwood	Parry Sound	0.7	3.5	30
	Tobermory	Parry Sound	2.9	4.1	36
	Harrisville	Harbor Beach	-0.6	3.9	32
	Essexville	Lakeport	-1.0	10.0	90
	Harbor Beach	Harrisville	0.6	3.9	32
	Goderich	Harbor Beach	0.2	3.9	33
	Lakeport	Harbor Beach	0.2	4.6	47

Table A. Best Substitute Station Error Statistics (cm)

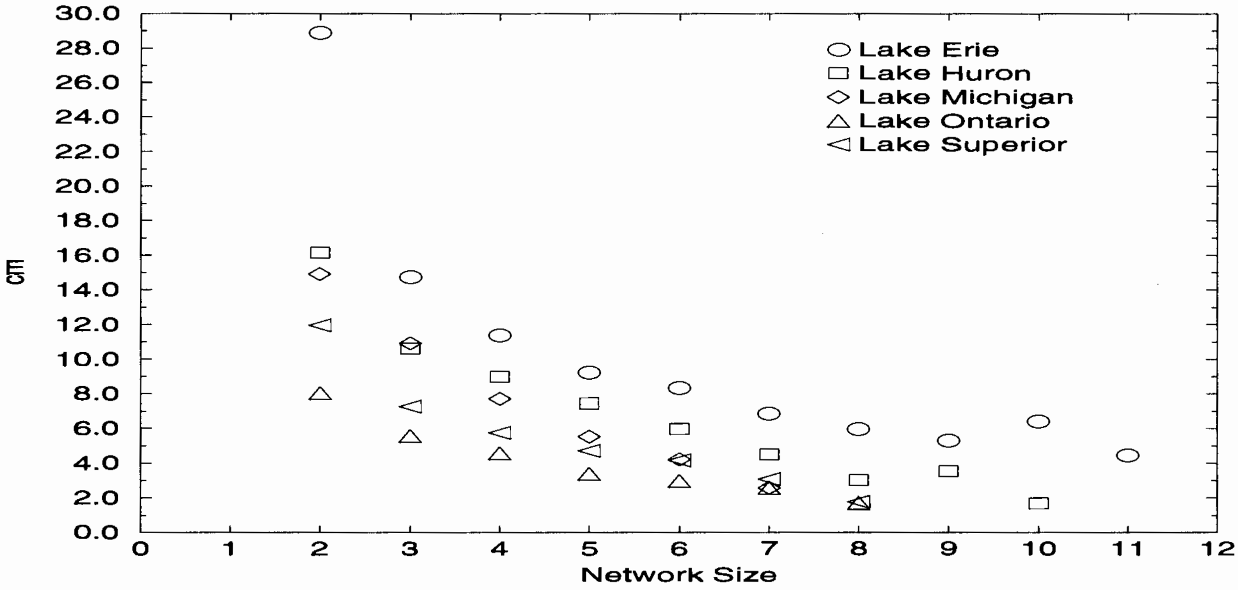
	Station	Substitute	Mean Error	Standard Error	Maximum Error
Lake Michigan	Mackinaw City	Port Inland	0.3	5.0	27
	Port Inland	Mackinaw City	-0.3	5.0	27
	Green Bay	Sturgeon Bay	0.7	11.0	73
	Sturgeon Bay	Kewaunee	0.2	4.0	27
	Kewaunee	Sturgeon Bay	-0.2	4.0	27
	Ludington	Sturgeon Bay	1.0	4.6	26
	Milwaukee	Holland	-0.2	4.6	49
	Holland	Milwaukee	0.2	4.6	49
	Calumet Harbor	Milwaukee	-0.1	7.0	49
Lake Ontario	Cape Vincent	Kingston	-1.0	1.7	12
	Kingston	Cape Vincent	1.0	1.7	12
	Oswego	Rochester	0.5	3.5	30
	Rochester	Olcott	-0.3	3.2	21
	Cobourg	Port Weller	-0.9	3.2	33
	Olcott	Port Weller	-0.3	2.1	15
	Toronto	Port Weller	0.0	2.5	16
	Port Weller	Olcott	0.3	2.1	15
	Burlington	Port Weller	-0.4	2.6	20
Lake Superior	Gros Cap	Point Iroquois	-1.0	1.7	14
	Point Iroquois	Gros Cap	1.0	1.7	14
	Michipicoten	Rossport	1.0	5.2	30
	Marquette	Ontonagon	0.4	5.0	27
	Rossport	Thunder Bay	-1.3	3.7	20
	Thunder Bay	Rossport	1.3	3.7	20
	Ontonagon	Grand Marais	-0.5	3.5	23
	Grand Marais	Ontonagon	0.5	3.5	23
	Duluth	Grand Marais	-0.7	4.7	30

Table B. Lowest Standard Error Subnetwork Configurations

	4-Station Network	6-Station Network	8-Station Network	
Lake Erie	Port Colborne	Buffalo Harbor	Buffalo Harbor	Fairport
	Erie	Port Dover	Sturgeon Point	Erieau
	Cleveland	Erie	Port Dover	Marblehead
	Kingsville	Fairport	Port Stanley	Bar Point
		Marblehead		
		Kingsville		
Lake Huron	Thessalon	Mackinaw City	Mackinaw City	Tobermory
	Tobermory	De Tour Village	Thessalon	Harrisville
	Harbor Beach	Little Current	Little Current	Goderich
	Goderich	Collingwood	Collingwood	Lakeport
		Goderich		
		Lakeport		
Lake Michigan	Mackinaw City	Mackinaw City		
	Kewaunee	Port Inland		
	Ludington	Kewaunee		
	Calumet Harbor	Ludington		
		Milwaukee		
		Calumet Harbor		
Lake Ontario	Cape Vincent	Cape Vincent	Cape Vincent	Cobourg
	Rochester	Oswego	Kingston	Toronto
	Toronto	Cobourg	Oswego	Port Weller
	Burlington	Olcott	Rochester	Burlington
		Toronto		
		Burlington		
Lake Superior	Gros Cap	Port Iroquois	Gros Cap	Thunder Bay
	Marquette	Michipicoten	Port Iroquois	Ontonagon
	Rossport	Marquette	Michipicoten	Grand Marais
	Duluth	Rossport	Marquette	Duluth
		Thunder Bay		
		Duluth		



a)



b)

Figure 33. Lowest first and second EOF mode a) standard error and b) maximum error for hourly water levels.

6. DISCUSSION

6.1. CORRELATIONS BETWEEN STATIONS

The cross correlation matrix formed in solving for EOF modes with hourly data shows how closely the signals at any two stations resemble each other. A correlation coefficient of 1.0 indicates complete correlation; 0.0 indicates no correlation; and -1.0 indicates complete correlation but 180 degrees out of phase. In Figures 34 to 37, station pairs for Lakes Erie, Huron, Michigan, and Superior with correlation coefficients greater than 0.90 are connected. For Lake Ontario, all stations are correlated with each other at a level greater than 0.95, so only correlation coefficients greater than 0.99 are shown (Figure 38). It can be seen that two stations that are close together generally have extremely high correlation coefficients.

The high correlations between nearby stations makes it natural to consider how large an error would result if the hourly water levels at one station were used as a direct substitute for hourly water levels at another station. Then, if a station is to be removed, data from a substitute station with a statistical error range can be used to estimate water level at the removed station. Statistics for the differences between pairs of stations were examined and for each station, the best substitute station was chosen based on the least standard error. The mean error, standard error, and maximum error for each substitute station are shown in Table A in the Appendix. The mean error may be attributed to long-term tectonic movement and/or gauge offset errors. Standard errors range from 1.7 cm for the Gros Cap - Port Iroquois pair on Lake Superior to 11.0 cm for the Green Bay - Sturgeon Bay pair on Lake Michigan. Maximum errors range from 12 cm for the Cape Vincent - Kingston pair on Lake Ontario to 90 cm for the Essexville - Lakeport pair on Lake Huron.

6.2. RECONSTRUCTION OF TIME SERIES AT INDIVIDUAL STATIONS

Since most of the hourly water level variance at any station on the Great Lakes is a linear combination of the first EOF mode and the second EOF mode (and the third EOF mode for Lake Huron), it should be possible to reconstruct most of the signal at any station based on EOF modes derived from subnetworks not containing that station. As an example, the best four, six, and eight station networks listed in Table B in the Appendix were used to obtain first and second EOF modes (and the third EOF mode for Lake Huron). The modes were scaled using eigenvectors obtained in EOF analyses of the complete networks to approximate the time series at stations not in the subnetworks. (For example, the water levels at Port Colborne, Erie, Cleveland, and Kingsville were used to reconstruct the time series at the rest of the Lake Erie stations.)

The error statistics for the difference between the reconstructed time series and the observed time series are shown in Table 5 for each of the subnetworks. In many cases, smaller standard errors were obtained for the reconstructed signals than for a single station substitution. However for other stations (marked in italics), substituting one single station resulted in a smaller standard error than attempting to reconstruct the signal from the EOF modes (compare Table 5 with Table A of the Appendix). This is true for station pairs close together which are highly correlated.

Table 5. Reconstructed Time Series Error Statistics (cm)

		4-Station Network		6-Station Network		8-Station Network	
Lake	**Station**	**Standard Error**	**Maximum Error**	**Standard Error**	**Maximum Error**	**Standard Error**	**Maximum Error**
Erie	Buffalo	*4.4*	*57*				
	Sturgeon Point	*2.9*	*38*	*2.7*	*35*		
	Port Colborne			*2.9*	*35*	2.6	31
	Port Dover	3.4	40				
	Erie					3.5	37
	Port Stanley	5.1	46	5.1	45		
	Fairport	3.7	51				
	Erieau	3.3	33	3.4	32		
	Cleveland			*4.3*	*74*	*4.8*	*81*
	Marblehead	4.1	48				
	Kingsville					3.1	34
	Bar Point	*5.7*	*88*	*5.7*	*77*		
Huron	Mackinaw	*4.2*	*38*				
	De Tour	2.0	16			1.9	13
	Thessalon			*2.6*	*21*		
	Little Current	4.3	35				
	Parry Sound	*4.0*	*36*	3.2	25	2.8	27
	Collingwood	*4.3*	*40*				
	Tobermory			2.3	19		
	Harrisville	2.8	33	3.0	36		
	Harbor Beach			2.9	29	2.5	26
	Lakeport	4.2	41				
Michigan	Port Inland	4.2	25				
	Sturgeon Bay	3.4	18	3.1	21		
	Milwaukee	4.2	34				
	Holland	4.0	33	3.7	34		
Ontario	Kingston	*1.8*	*15*	*2.1*	*16*		
	Oswego	2.9	25				
	Rochester			2.5	19		
	Cobourg	3.0	29				
	Olcott	*2.2*	*12*			1.9	11
	Port Weller	1.7	10	1.5	9		

Table 5. Reconstructed Time Series Error Statistics (cm)

		4-Station Network		6-Station Network		8-Station Network	
Lake	**Station**	**Standard Error**	**Maximum Error**	**Standard Error**	**Maximum Error**	**Standard Error**	**Maximum Error**
Superior	Gros Cap			*2.7*	*22*		
	Point Iroquois	*2.4*	*21*				
	Michipicoten	3.8	28				
	Rossport					3.3	15
	Thunder Bay	3.6	23				
	Ontonagon	2.6	16	2.5	13		
	Grand Marais	3.1	27	2.8	28		

6.3. LAKE HURON - LAKE MICHIGAN SYSTEM

It has been noted that the higher frequency variability (1-5 day periods) of the Lake Huron and Lake Michigan mean lake levels is greater than those of the other three lakes. This is because the two lakes are connected at the Straits of Mackinac and water can be transferred back and forth between the lakes by meteorological forcing. Since the two lakes have nearly identical surface areas, if the mean lake levels of the two lakes are averaged, the effect of the transfer of water between lakes should be eliminated. The first EOF modes of the 2-day averaged data for Lakes Huron and Michigan are shown in Figure 39 together with their mean (offset for comparison). The mean of the two lake levels is a much smoother time series, similar to the mean lake levels of the other lakes. The mean of the hourly first EOF modes for Lakes Huron and Michigan (Figure 40) also shows a significant reduction in higher frequency energy compared to the first EOF mode for each individual lake. Therefore, calculation of the mean of the Lake Huron and Lake Michigan first EOF modes is necessary to eliminate the meteorological forcing effects.

6.4. STRATEGY FOR EVALUATING NETWORK SIZE REDUCTIONS

The results of this investigation can provide guidelines for evaluating the effects of network size reductions. First, a desired level of accuracy should be chosen for one or more lake level statistics, based on users requirements. Standard errors and maximum errors could be specified for 2-day averaged mean lake levels. Standard errors and maximum errors could also be specified for hourly mean lake levels and lake level tilts and/or the total signal at any particular station that may be eliminated.

Then, an examination of Figures 22 and 33 can give an idea of the minimum network size that would be needed for each lake to produce the desired accuracy. Figure 22 shows subnetwork errors for the first EOF mode for 2-day averaged data; Figure 33 shows subnetwork errors for the first and second EOF modes for hourly data. The larger the acceptable error, the smaller the network can be. Once the reduced network size is decided, the lowest error subnetwork can be chosen as the new network. However, there will be many other combinations which will be almost as good as the lowest error subnetwork, so other considerations can be taken into account in deciding which stations could be removed.

The maps in Figures 34 to 38 showing the highest correlation coefficients between stations are helpful in identifying station pairs that give almost identical water level information. This often occurs where a U. S. and Canadian station are located very close together (e.g. Fermi and Bar Point, Buffalo Harbor and Port Colborne, De Tour Village and Thessalon, Point Iroquois and Gros Cap, Olcott and Port Weller, Cape Vincent and Kingston). Other stations that have high correlation coefficients with more than one nearby stations could also be considered duplicative (e.g. Sturgeon Point, Kingsville, Parry Sound, or any station on Lake Ontario). When a smaller network has been chosen, an EOF analysis can be carried out and the resulting modes can be compared with the EOF modes from the full network to see if the desired accuracy levels have been met.

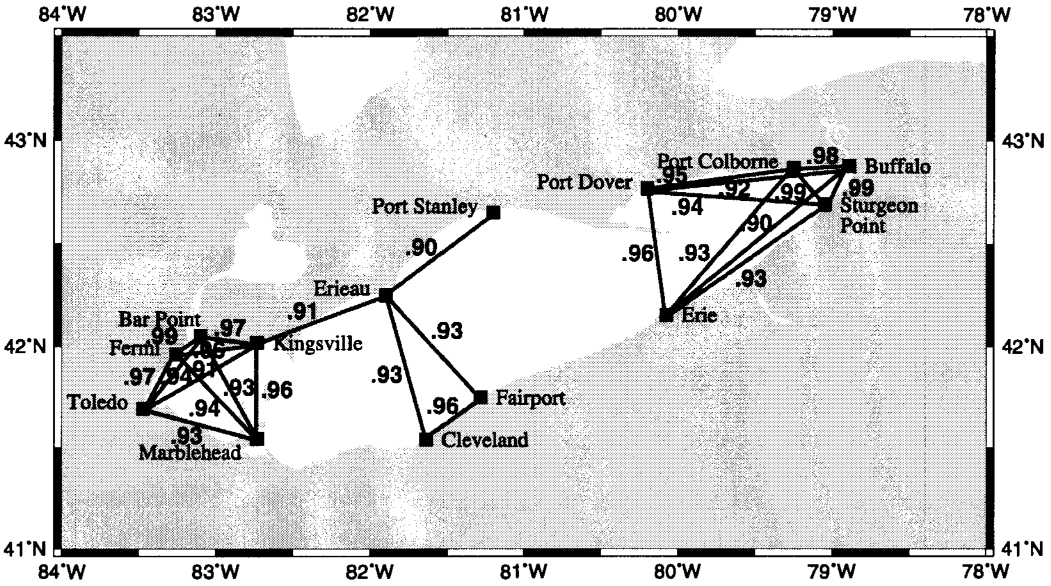


Figure 34. Correlation coefficents greater than 0.90 for Lake Erie 1990 hourly water levels.

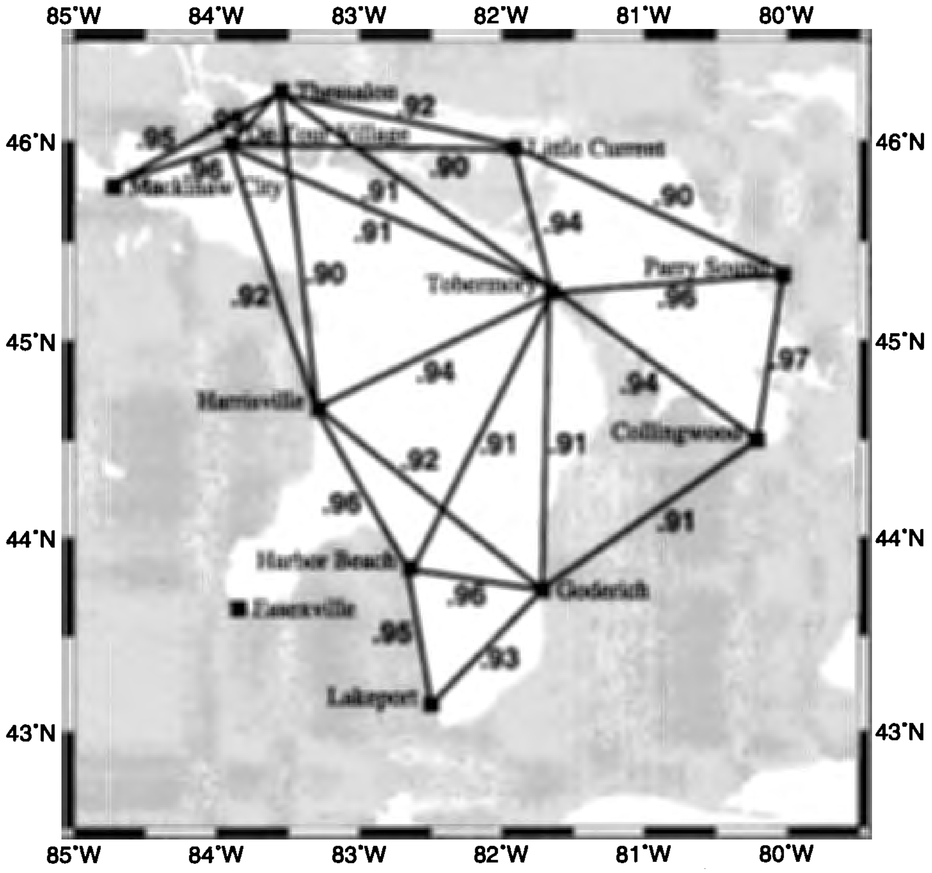


Figure 35. Correlation coefficients greater than 0.90 for Lake Huron 1990 hourly water levels.

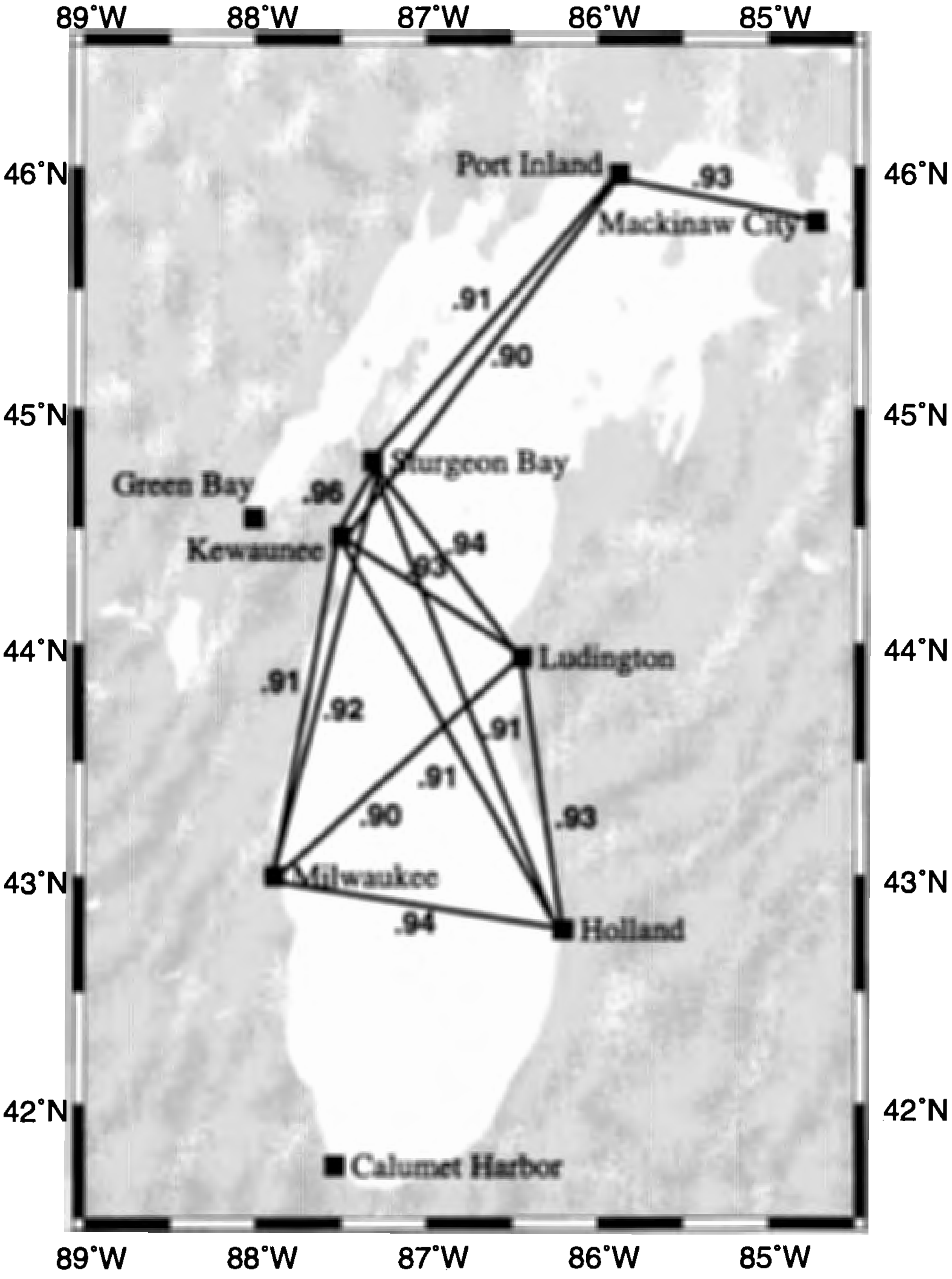


Figure 36. Correlation coefficients greater than 0.90 for Lake Michigan 1990 hourly water levels.

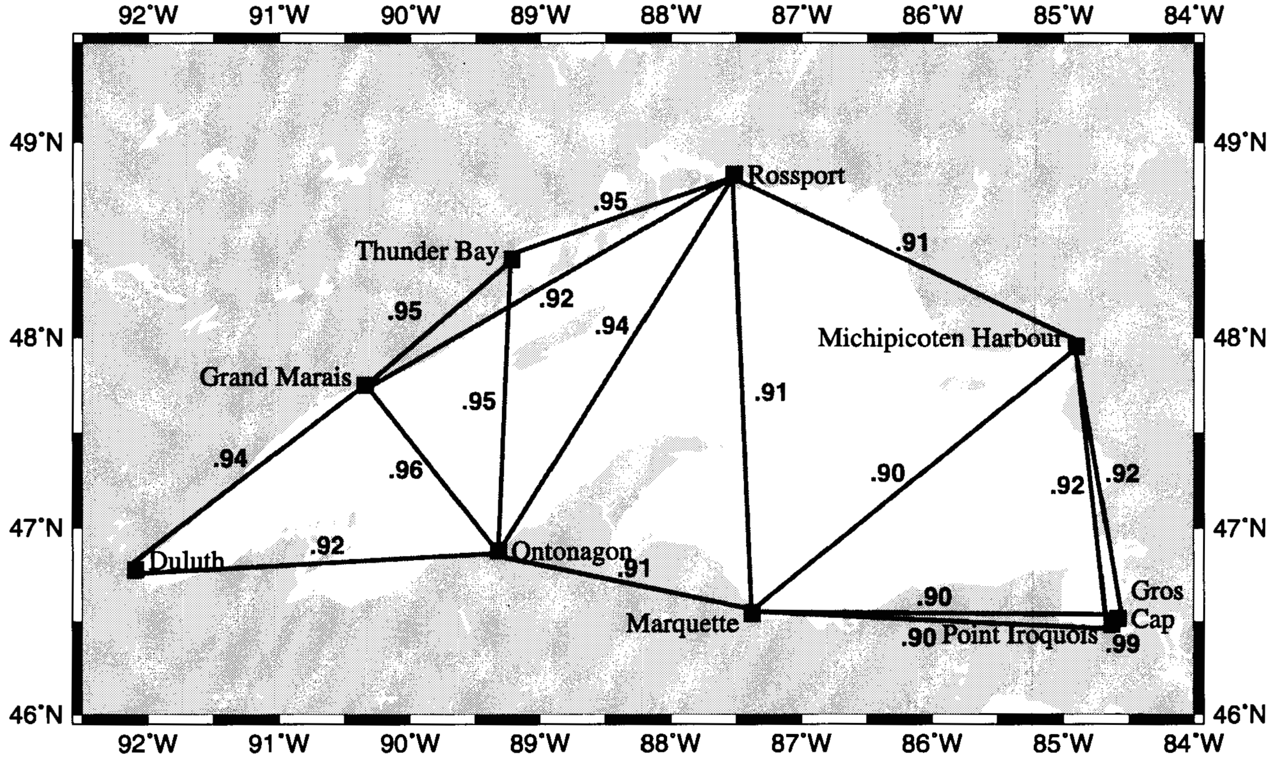


Figure 37. Correlation coefficients greater than 0.90 for Lake Superior 1990 hourly water levels.

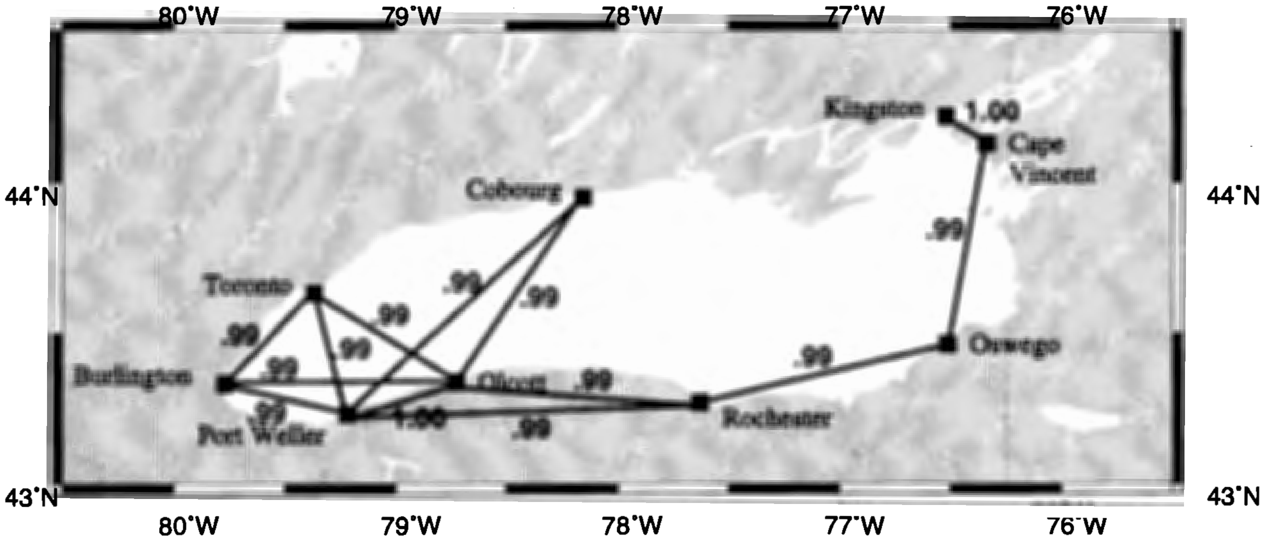


Figure 38. Correlation coefficients greater than 0.99 for Lake Ontario 1990 hourly water levels.

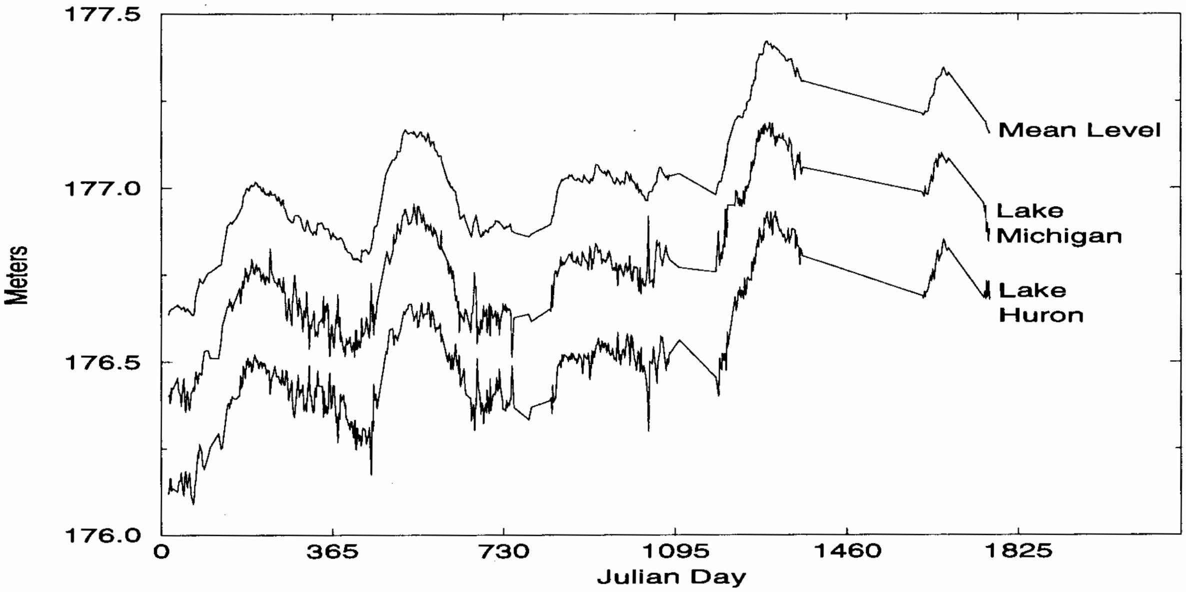


Figure 39. First EOF modes for Lake Michigan and Lake Huron and their mean for 2-day averaged water levels for 1990-1994 (offset for comparison).

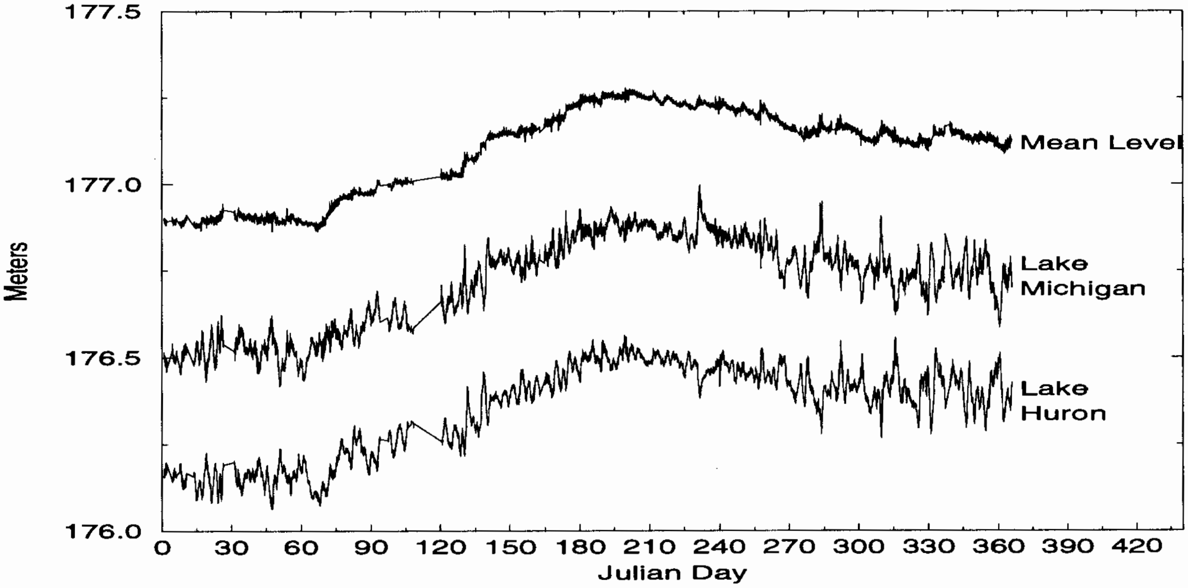


Figure 40. First EOF modes for Lake Michigan and Lake Huron and their mean for hourly water levels for 1990 (offset for comparison).