

LAKE-LEVEL FREQUENCY ANALYSIS FOR DEVILS LAKE, NORTH DAKOTA

By Gregg J. Wiche and Aldo V. Vecchia

U.S. GEOLOGICAL SURVEY

Open-File Report 95-123

Prepared in cooperation with the
NORTH DAKOTA STATE WATER COMMISSION

Bismarck, North Dakota

1995



**U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary**

**U.S. GEOLOGICAL SURVEY
GORDON P. EATON, Director**

Any use of trade, product, or firm names is for descriptive purposes only
and does not imply endorsement by the U.S. Government.

For additional information
write to:

District Chief
U.S. Geological Survey
Water Resources Division
821 East Interstate Avenue
Bismarck, ND 58501

Copies of this report can be
purchased from:

U.S. Geological Survey
Earth Science Information Center
Open-File Reports Section
Box 25286, MS 517
Denver Federal Center
Denver, CO 80225

CONTENTS

Abstract	1
Introduction	2
Description of study area	2
Lake-level fluctuations	4
Previous investigations	5
Approach	7
Lake-level frequency analysis computed using an annual lake-volume model	8
Computation of annual lake-volume data	8
Annual lake-volume model	10
Analysis of model output	17
Parameter-uncertainty analysis	21
Lake-level frequency analysis computed using a water mass-balance model	21
Computation of monthly precipitation, evaporation, and inflow data	21
Statistical tests for trends and long-range dependence	28
Water mass-balance model	31
Analysis of model output	36
Parameter-uncertainty analysis	42
Reduction of parameter uncertainty in seasonal inflow distributions	44
Summary and conclusions	48
References	53
Supplement 1. Parameter-uncertainty distributions and model fitting procedure	56
Parameter-uncertainty distribution for the annual lake-volume model	56
Model fitting procedure for quarterly precipitation, evaporation, and inflow	58
Parameter-uncertainty distribution for the water mass-balance model	60
Variance of adjusted means of seasonal inflow distributions	63

FIGURES

1. Map showing location of the Devils Lake Basin	3
2. Graph showing historic water level for Devils Lake, 1867-1994	6
3. Graph showing annual mean lake volume for year t for Devils Lake, 1867-1993	9
4. Graph showing difference between annual maximum lake volume for year t for Devils Lake and annual mean lake volume for year t , 1901-93	9
5. Graphs showing standardized residuals, $W^*(t)$, from autoregressive moving average model of annual mean lake volumes for Devils Lake, 1867-1993, and normal probability plot	12
6. Graphs showing standardized residuals, $Z^*(t)$, from autoregressive moving average model of differenced annual lake volumes for Devils Lake, 1901-93, and normal probability plot	14
7. Graph showing relation of standardized residuals, $W^*(t)$, from autoregressive moving average model of annual mean lake volumes for Devils Lake to standardized residuals, $Z^*(t)$, from autoregressive moving average model of differenced annual lake volumes	15
8. Graphs showing standardized residuals, $R^*(t)$, from autoregressive moving average model of differenced annual lake volumes for Devils Lake after nonlinear adjustment, 1901-93, and normal probability plot	16
9. Graphs showing unconditional simulation lake-level trace from the annual lake-volume model and corresponding lake-level changes for Devils Lake	18
10. Graph showing exceedance levels for Devils Lake computed from 2,000 50-year traces from the annual lake-volume model using fixed parameters and parameters adjusted for uncertainty	20
11. Graph showing annual precipitation for the city of Devils Lake, 1950-93	23
12. Graph showing annual lake evaporation for Devils Lake estimated from unadjusted pan-evaporation data for Mandan and Dickinson, 1950-93	26

FIGURES, Continued

13. Graph showing annual lake evaporation for Devils Lake estimated from adjusted pan- evaporation data for Mandan and Dickinson, 1950-93.....	26
14. Graph showing relation of Class A pan evaporation at Mandan and Dickinson to Class A pan evaporation at Devils Lake, 1951-71	28
15. Graph showing estimated annual inflow to Devils Lake, 1950-93	29
16. Graphs showing autocorrelation functions for annual precipitation, annual evaporation, and natural logarithm of annual inflow for Devils Lake	30
17. Graphs showing quarterly precipitation, evaporation, and inflow for Devils Lake, 1950-93	32
18. Graph showing modeled and recorded quarterly storage changes for Devils Lake, 1950-93	33
19. Graphs showing relation of summer inflow to fall inflow for Devils Lake, 1950-93.....	39
20. Graphs showing unconditional simulation lake-level trace from the water mass-balance model and corresponding lake-level changes for Devils Lake	41
21. Graph showing exceedance levels for Devils Lake computed from 2,000 50-year traces from the water mass-balance model using fixed parameters and parameters adjusted for uncertainty.....	43
22. Graph showing spring streamflow for the Red River of the North at Grand Forks, 1883-1993	45
23. Graph showing relation of spring streamflow for the Red River of the North at Grand Forks to spring inflow for Devils Lake, 1950-93.....	46
24. Graph showing exceedance levels for Devils Lake computed from 2,000 50-year traces from the water mass-balance model using parameters adjusted for uncertainty and reduced on the basis of streamflow for the Red River of the North at Grand Forks	49
25. Graph showing exceedance levels for Devils Lake computed from 2,000 50-year traces from the water mass-balance model with each trace conditioned only on the starting lake level.....	50

TABLES

1. Largest lake-level rises and declines for Devils Lake for 1950-93	5
2. Selected statistics of recorded lake-level changes for 1901-93 and corresponding values generated from the annual lake-volume model.....	19
3. Exceedance levels for Devils Lake computed from 2,000 50-year traces from the annual lake-volume model and the water mass-balance model	22
4. Mean monthly percentage of annual lake evaporation and mean monthly lake evaporation for Devils Lake	24
5. Selected statistics for monthly Class A pan-evaporation data for May through September for Mandan and Dickinson	27
6. Selected statistics of trend tests for annual precipitation, evaporation, and inflow for Devils Lake, 1950-93.....	29
7. Selected statistics of precipitation, evaporation, and inflow data for 1950-93 and corresponding values generated from the water mass-balance model.....	37
8. Selected statistics of recorded lake-level changes for 1901-93 and corresponding values generated from the water mass-balance model.....	40
9. Estimated means of transformed quarterly inflows before and after adjustment to reduce parameter uncertainty	48

CONVERSION FACTORS AND VERTICAL DATUM

Multiply	By	To obtain
acre	4,047	hectare
acre-foot	1,233	cubic meter
cubic foot per second	0.02832	cubic meter per second
foot	0.3048	meter
inch	25.4	millimeter
mile	1.609	kilometer
square mile	2.590	square kilometer

Sea level: In this report, "sea level" refers to the National Geodetic Vertical Datum of 1929 (NGVD of 1929)--a geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, formerly called Sea Level Datum of 1929.

Lake-Level Frequency Analysis for Devils Lake, North Dakota

By Gregg J. Wiche *and* Aldo V. Vecchia

Abstract

Two approaches were used to estimate future lake-level probabilities for Devils Lake. The first approach is based on an annual lake-volume model, and the second approach is based on a statistical water mass-balance model that generates seasonal lake volumes on the basis of seasonal precipitation, evaporation, and inflow.

Autoregressive moving average models were used to model the annual mean lake volume and the difference between the annual maximum lake volume and the annual mean lake volume. Residuals from both models were determined to be uncorrelated with zero mean and constant variance. However, a nonlinear relation between the residuals of the two models was included in the final annual lake-volume model.

Because of high autocorrelation in the annual lake levels of Devils Lake, the annual lake-volume model was verified using annual lake-level changes. The annual lake-volume model closely reproduced the statistics of the recorded lake-level changes for 1901-93 except for the skewness coefficient. However, the model output is less skewed than the data indicate because of some unrealistically large lake-level declines.

The statistical water mass-balance model requires as inputs seasonal precipitation, evaporation, and inflow data for Devils Lake. Analysis of annual precipitation, evaporation, and inflow data for 1950-93 revealed no significant trends or long-range dependence so the input time series were assumed to be stationary and short-range dependent.

Normality transformations were used to approximately maintain the marginal probability distributions; and a multivariate, periodic autoregressive model was used to reproduce the correlation structure. Each of the coefficients in the model is significantly different from zero at the 5-percent significance level. Coefficients relating spring inflow from one year to spring and fall inflows from the previous year had the largest effect on the lake-level frequency analysis.

Inclusion of parameter uncertainty in the model for generating precipitation, evaporation, and inflow indicates that the upper lake-level exceedance levels from the water mass-balance model are particularly sensitive to parameter uncertainty. The sensitivity in the upper exceedance levels was caused almost entirely by uncertainty in the fitted probability distributions of the quarterly inflows. A method was developed for using long-term streamflow data for the Red River of the North at Grand Forks to reduce the variance in the estimated mean.

Comparison of the annual lake-volume model and the water mass-balance model indicates the upper exceedance levels of the water mass-balance model increase much more rapidly than those of the annual lake-volume model. As an example, for simulation year 5, the 99-percent

exceedance for the lake level is 1,417.6 feet above sea level for the annual lake-volume model and 1,423.2 feet above sea level for the water mass-balance model. The rapid increase is caused largely by the record precipitation and inflow in the summer and fall of 1993. Because the water mass-balance model produces lake-level traces that closely match the hydrology of Devils Lake, the water mass-balance model is superior to the annual lake-volume model for computing exceedance levels for the 50-year planning horizon.

INTRODUCTION

Devils Lake Basin in northeastern North Dakota is a 3,810-square-mile closed basin in the drainage of the Red River of the North (fig. 1). About 3,320 square miles of the total 3,810 square miles is tributary to Devils Lake; the remainder is tributary to Stump Lake. The Devils Lake area has been a popular recreational area during the last 110 years, and Devils Lake has been the main attraction for much of the recreation. Unlike most terminal lakes, Devils Lake has been a productive sport fishing lake intermittently since settlers arrived in the early 1880's. Ideas and plans to stabilize the lake level to protect recreation at Devils Lake were first mentioned more than 50 years ago. In the late 1930's and early 1940's, various plans were developed by the U.S. Army Corps of Engineers and the Bureau of Reclamation to divert water from the Missouri River to Devils Lake.

In response to rising lake levels from about 1969 through the early 1980's, the U.S. Army Corps of Engineers (1984) completed a feasibility study of possible flood-control projects to protect cities, roads, and other properties around Devils Lake. The U.S. Army Corps of Engineers also conducted a reconnaissance study for lake-level stabilization of Devils Lake, but the study will not be approved until additional analyses of future lake-level probabilities and associated economic damage estimates are completed. An understanding of the hydrology of Devils Lake and better knowledge of future lake-level probabilities is needed as a basis for implementation of flood-control or lake-stabilization projects. The U.S. Geological Survey, in cooperation with the North Dakota State Water Commission, conducted a lake-level frequency study in an effort to estimate future lake-level probabilities. Specific objectives of the study were to (1) estimate lake levels for Devils Lake using an analysis of historic lake levels and (2) estimate lake levels for Devils Lake using a statistical water mass-balance model. This report describes the results of the study. Lake-level probability data will be used by the U.S. Army Corps of Engineers to develop damage-frequency relations for Devils Lake. Data used to develop time-series models will be limited to hydrologic and meteorologic data available through June 1994.

Description of Study Area

The topographic relief and surficial landforms in the Devils Lake Basin are of glacial origin. A large number of shallow depressions and potholes occur throughout the basin. Many of these depressions are connected by poorly defined channels and swales.

The eastern, western, and northern boundaries of the Devils Lake Basin are poorly defined low divides. The southern boundary is a series of recessional moraines that lie between Devils Lake and the Sheyenne River. The major subbasins in the Devils Lake Basin and the principal streams draining the subbasins are shown in figure 1. Edmore, Starkweather, and Calio Coulees originate in southern Cavalier County and flow in a south-southwesterly direction. Mauvais Coulee originates along the southern flanks of the Turtle Mountains 300 to 400 feet above the elevation of Devils Lake and flows in a southerly direction.

Before 1979, streamflow from the tributaries flowed into the interconnected chain of lakes (Sweetwater Lake, Morrison Lake, Dry Lake, Mikes Lake, Chain Lake, Lake Alice, and Lake Irvine), and all streamflow from the chain of lakes flowed downstream through Big Coulee into Devils Lake (fig. 1). In

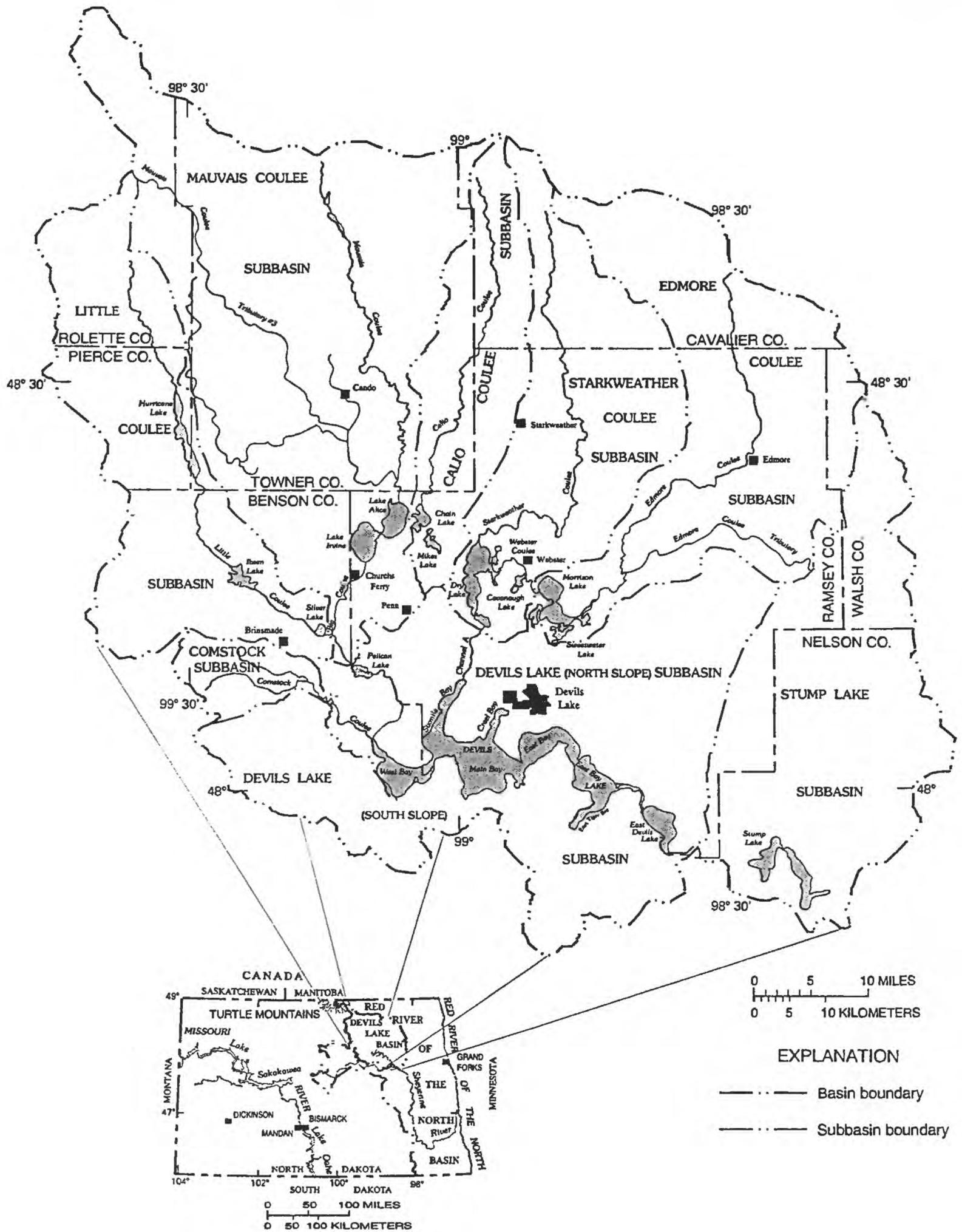


Figure 1. Location of the Devils Lake Basin.

1979, the Ramsey County and Cavalier County Water Management Boards constructed Channel A, which connects Dry Lake to Sixmile Bay on Devils Lake (fig. 1). A levee was constructed across the natural outlet of Dry Lake in 1979. The construction of Channel A and the levee on Dry Lake modified the drainage pattern in the basin. Discharge from Dry Lake to Sixmile Bay via Channel A is regulated by an adjustable head-gate control at the south shore of the lake. Runoff into Sweetwater, Morrison, and Dry Lakes discharges through Channel A into Devils Lake; the remaining runoff discharges along the natural watercourse through Big Coulee into Devils Lake. A small quantity of runoff also enters Devils Lake by overland flow and in small tributaries from drainage areas adjacent to the lake.

Lake-Level Fluctuations

Since glaciation, the lake level of Devils Lake has fluctuated from about 1,454 feet above sea level, the natural spill elevation of the lake, to about 1,400 feet above sea level (Aronow, 1957). According to Bluemle (1981), the elevation of Devils Lake was more than 1,440 feet above sea level 8,500 years before present. Callender (1968, p. 261) made various chemical analyses of sediment samples from Devils Lake to provide a lake-level chronology for the past 6,500 years. Callender (1968) concluded that

"The lake was dry during the last part of the Hypsithermal (6,500 years before present) interval. The level rose and then declined several times between 6,000 and 2,500 years before present, after which a peat was deposited in Creel Bay approximately 1,340 years ago. Several more lake-level fluctuations culminated in a very saline, low-water stage 500 years before present, when oak trees grew on the dry surface sediment of East Stump Lake. The level subsequently rose until 1800 A.D., declined to a low-water stage in 1940 A.D., rose until 1951 A.D., and steadily declined from that time to the present [1968]. Comparison of the Devils Lake chronology with those from other regions indicates that major climatic changes which caused significant fluctuations in the lake level may have extended beyond the northern Great Plains region."

Aronow (1955, 1957) analyzed abandoned shorelines, water-deposited sand and gravel deposits containing buried soils and vertebrate remains, and rooted stumps uncovered by receding water around Stump Lake. In general, Aronow's interpretation (1955, 1957) of lake-level fluctuations is similar to Callender's interpretation (1968) although some differences do exist. Aronow (1955, 1957) indicated that a lowering of lake levels in the Devils Lake Basin occurred during a dry period in the 15th and 16th centuries, as evidenced by the growth of burr oak in Stump Lake. According to Brooks (1951), this dry period occurred throughout most of western North America. Following this dry period, there was a general rise in lake levels from the mid-1500's until the mid- to late 1800's. This period of rising lake levels commonly is referred to as the Little Ice Age. In a more recent study, Bluemle (1988) used radiocarbon dates of soils and concluded that Devils Lake overflowed into Stump Lake in the last 1,800 years.

All of these studies indicate that large and frequent lake-level fluctuations of 20 to 40 feet occur every few hundred years. A rising or declining lake level seems to be a more normal condition for Devils Lake than a stable lake level.

U.S. Geological Survey personnel have a long history of documenting lake-level fluctuations of Devils Lake. Upham (1895, p. 595) indicated that the lake level of Devils Lake was 1,441 feet above sea level in 1830. He based this lake level on a large, dense stand of timber that grew at and above 1,441 feet above sea level. Below 1,441 feet above sea level, only scattered trees and brush existed. Captain H.H. Heerman informed Upham that, based on tree-ring chronology, the largest tree cut below 1,441 feet above sea level was 57 years old in 1887. Thus, Upham (1895, p. 595) concluded that in 1830 (57 years before 1887) the lake level of Devils Lake was 1,441 feet above sea level. No documented records of lake levels are available before 1867.

Lake levels of Devils Lake were recorded sporadically from 1867 to 1901 (fig. 2). In 1901, the U.S. Geological Survey established a gage at Devils Lake. For the period of record at Devils Lake, the maximum lake level occurred in 1867; the lake level was 1,438 feet above sea level and the lake had a surface area of about 140 square miles. From 1867 to 1940, the lake level declined almost continuously until it reached a recorded low of 1,400.9 feet above sea level and the lake was a shallow brackish body of water that had a surface area of about 10.2 square miles. From 1940 to 1956, the lake level generally rose. From 1956 to 1968, the lake level generally declined. From 1968 to 1987, the lake level generally rose until it reached a peak of 1,428.8 feet above sea level, which was the highest lake level in almost 100 years. At the peak in 1987, Devils Lake had a surface area of about 94 square miles. From 1987 to 1992, the lake level of Devils Lake declined rapidly to 1,422.6 feet above sea level. Much greater than normal precipitation during June through August 1993 caused widespread flooding throughout the central and north-central United States, including in the Devils Lake Basin. Rainfall on Devils Lake, streamflow during the summer and fall of 1993, and snowmelt runoff in 1994 caused the lake level of Devils Lake to rise to 1,430.6 feet above sea level by June 30, 1994.

The largest lake-level rises and declines for Devils Lake for 1950-93 are shown in table 1. Devils Lake has undergone rapid changes in lake level in some years. The hydroclimatic conditions in the Devils Lake area constrain the largest lake-level declines to be grouped between 1.8 and 2.4 feet (table 1).

Table 1. Largest lake-level rises and declines for Devils Lake for 1950-93

Year	Lake-level rise (feet)	Year	Lake-level decline (feet)
1950	8.4	1976	2.4
1969	6.3	1952	2.0
1993	5.2	1972, 1988	1.9
1974	4.9	1959, 1961, 1963	1.8
1979	4.9	--	--

Previous Investigations

The U.S. Geological Survey, in cooperation with several State and Federal agencies, completed several studies that describe the hydrology of Devils Lake; however, none of the studies describe future lake-level probabilities. Swenson and Colby (1955) conducted a water-quality study to determine the dissolved-solids mass in Devils Lake. They also developed a water balance of Devils Lake and estimated how much inflow would be needed to maintain the lake at 1,425 feet above sea level. Mitten and others (1968) continued the water-quality study of Devils Lake that was started by Swenson and Colby.

Wiche (1986) reviewed the climatologic and hydrologic data to determine the causes of lake-level fluctuations of Devils Lake. He concluded that the lake level of Devils Lake fluctuates in response to climate variability, but the hydrologic characteristics of the Devils Lake Basin distort the hydrologic response. Wiche (1986) also conducted a literature search to determine what approaches have been used to estimate future lake-level probabilities of terminal lakes.

Wiche and others (1986) completed a report at the request of the U.S. Justice Department that was used in a lawsuit between 101 Ranch and the United States of America, Civil No. A2-81-89. They determined the exceedance probability for a given lake-level rise by computing the probability for a given inflow and then converting the inflow to a lake-level rise by using the lake-level capacity table for the

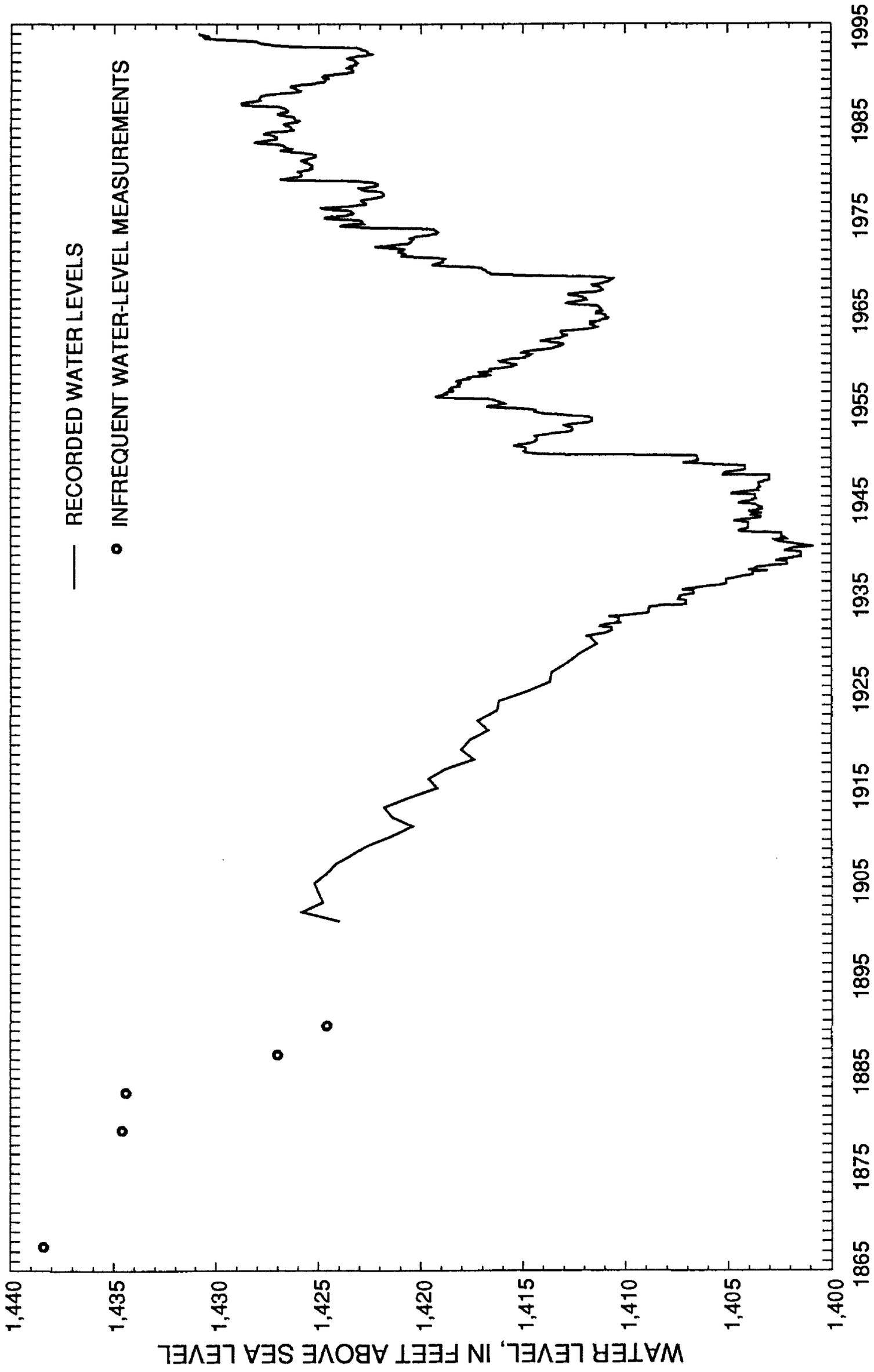


Figure 2. Historic water level for Devils Lake, 1867-1994.

desired starting lake level. The exceedance probability for a given inflow was computed by using the procedures outlined by the U.S. Water Resources Council (1981). The report was used to develop the list of "uncontroverted facts" agreed upon by attorneys for all parties.

Ryan and Wiche (1988) described the hydrology of the chain of lakes upstream of Devils Lake and the effect of the chain of lakes on lake-level fluctuations of Devils Lake. Ryan and Wiche (1988) also presented multilake management situations by simulating lake levels of Devils Lake for different storage conditions in the chain of lakes, for different runoff conditions in the Devils Lake Basin, and for different outlet options for Devils Lake. Various storage conditions in the chain of lakes, the different runoff conditions in the Devils Lake Basin, and the maximum outlet releases to the Sheyenne River were selected in consultation with the U.S. Army Corps of Engineers (Kent Pederson and Bill Spychalla, written commun., 1986).

Woodbury and Padmanabhan (1989) developed time-series models of annual average lake levels and annual maximum deviations from the historic average lake levels. Together, the models were used to develop a lake-level frequency relation for Devils Lake.

To estimate lake-level probabilities for the Great Salt Lake, James and others (1977) developed an approach that links time-series models of inflow and outflow data to a water mass-balance model. The primary criterion for calibrating the model was to preserve the mean, standard deviation, and cross-correlation matrices among the input data.

In 1986, the U.S. Geological Survey, in cooperation with the North Dakota State Water Commission, began a study to improve estimates of the water balance of Devils Lake. The U.S. Geological Survey lead an energy-budget and mass-transfer evaporation study (Wiche, 1992), and the North Dakota State Water Commission lead a study of ground-water movement into and out of Devils Lake (Pusc, 1993).

Approach

Many approaches have been used to estimate future lake-level probabilities of terminal lakes (Wiche, 1986), but no standard method currently exists. Two approaches were used in this study. The first approach was an analysis of historic lake levels of Devils Lake as outlined by the U.S. Army Corps of Engineers (written commun., June 11, 1993). The approach is based on an annual lake-volume (ALV) model and is similar to the approach used by Woodbury and Padmanabhan (1989) in a study of lake-level frequency relations for Devils Lake.

The second approach is based on a statistical water mass-balance (WMB) model that generates seasonal lake volumes on the basis of seasonal precipitation, evaporation, and inflow. A multivariate time-series model was developed to generate future realizations of seasonal precipitation, evaporation, and inflow, which are used in conjunction with the WMB model to generate future lake-level traces. The second approach is similar to the approach used by Bowles and James (1986) and James and others (1977). The approach also is similar to an approach outlined by Yevjevich (1982) under contract to the U.S. Army Corps of Engineers.

One advantage of the annual lake-volume approach (hereafter referred to as the ALV approach) is that estimates of precipitation, evaporation, and inflow are not required as inputs. Therefore, the ALV approach is easier to implement than the water mass-balance approach (hereafter referred to as the WMB approach). Another advantage of the ALV approach is that it is based on a longer period of record than the WMB approach because annual lake levels of Devils Lake are available from the late 1800's to the present (1994) but evaporation and inflow data are available only from 1950 to the present.

Despite relative ease of implementation, the ALV approach has several disadvantages. Because the ALV approach does not account specifically for precipitation, evaporation, inflow, and lake geometry, the lake-level traces produced are not always realistic on a hydrologic basis. Another disadvantage to the ALV

approach is that verification of lake-level probabilities produced by the model is difficult, largely because of extremely high autocorrelation from year to year (temporal persistence) in the recorded lake levels of Devils Lake (fig. 2). Therefore, model performance cannot be measured by how well the lake-level probabilities produced by the ALV model match the statistics for the recorded lake-level changes. The WMB model is easier to verify because it is based on precipitation, evaporation, and inflows to Devils Lake that are much less persistent than the lake levels.

The WMB approach has the distinct advantage of producing lake-level traces that agree with the water mass-balance equation for Devils Lake. Various inlet or outlet options also can be incorporated into the model in subsequent lake-level stabilization studies. The disadvantage to the WMB approach is that the model requires more extensive data collection and model development than the ALV approach.

LAKE-LEVEL FREQUENCY ANALYSIS COMPUTED USING AN ANNUAL LAKE-VOLUME MODEL

Computation of Annual Lake-Volume Data

Most of the available data for Devils Lake before 1950 consist of annual maximum and minimum lake levels. In order to develop a time-series model that uses these data, an approach similar to that used by Woodbury and Padmanabhan (1989) was used in this study. In this approach, the time series of annual maximum and minimum lake levels are converted to maximum and minimum lake volumes using the Devils Lake elevation-capacity table (Dale Frink, North Dakota State Water Commission, written commun., 1994). The maximum lake volume for year t is expressed as

$$V_{max}(t) = X(t) + D(t), \quad (1)$$

where

$V_{max}(t)$ is the maximum lake volume for year t , in thousands of acre-feet;

$X(t)$ is the mean lake volume for year t , in thousands of acre-feet; and

$D(t)$ is the difference between the maximum lake volume for year t and the mean lake volume for year t , in thousands of acre-feet.

$X(t)$ is expressed as

$$X(t) = \frac{[V_{max}(t) + V_{min}(t)]}{2}, \quad (2)$$

where

$V_{min}(t)$ is the minimum lake volume for year t , in thousands of acre-feet;

and $D(t)$ is expressed as

$$D(t) = \left[\frac{V_{max}(t) - V_{min}(t)}{2} \right]. \quad (3)$$

Recorded values of $X(t)$ are shown in figure 3, and recorded values of $D(t)$ are shown in figure 4.

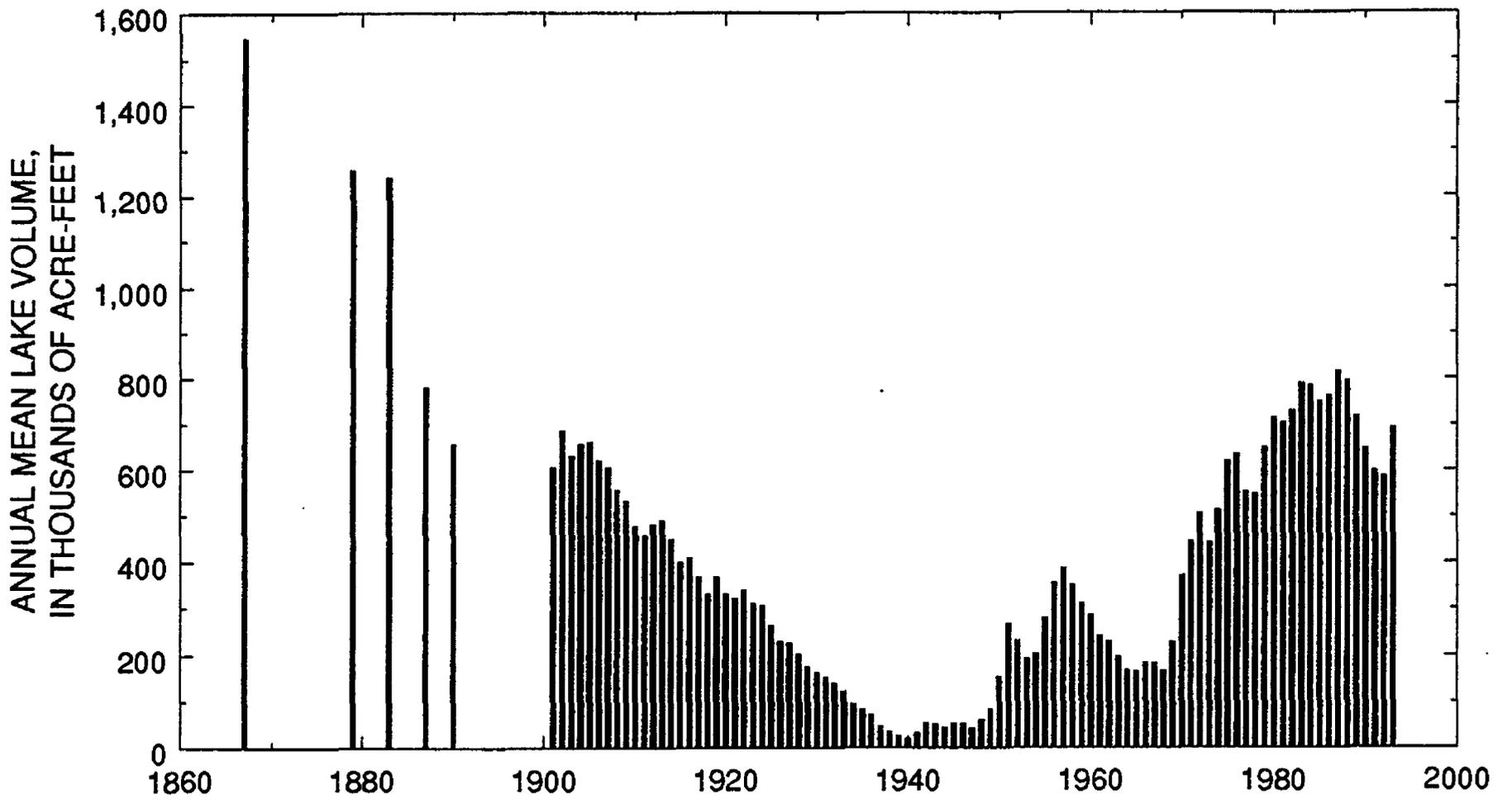


Figure 3. Annual mean lake volume for year t for Devils Lake, 1867-1993.

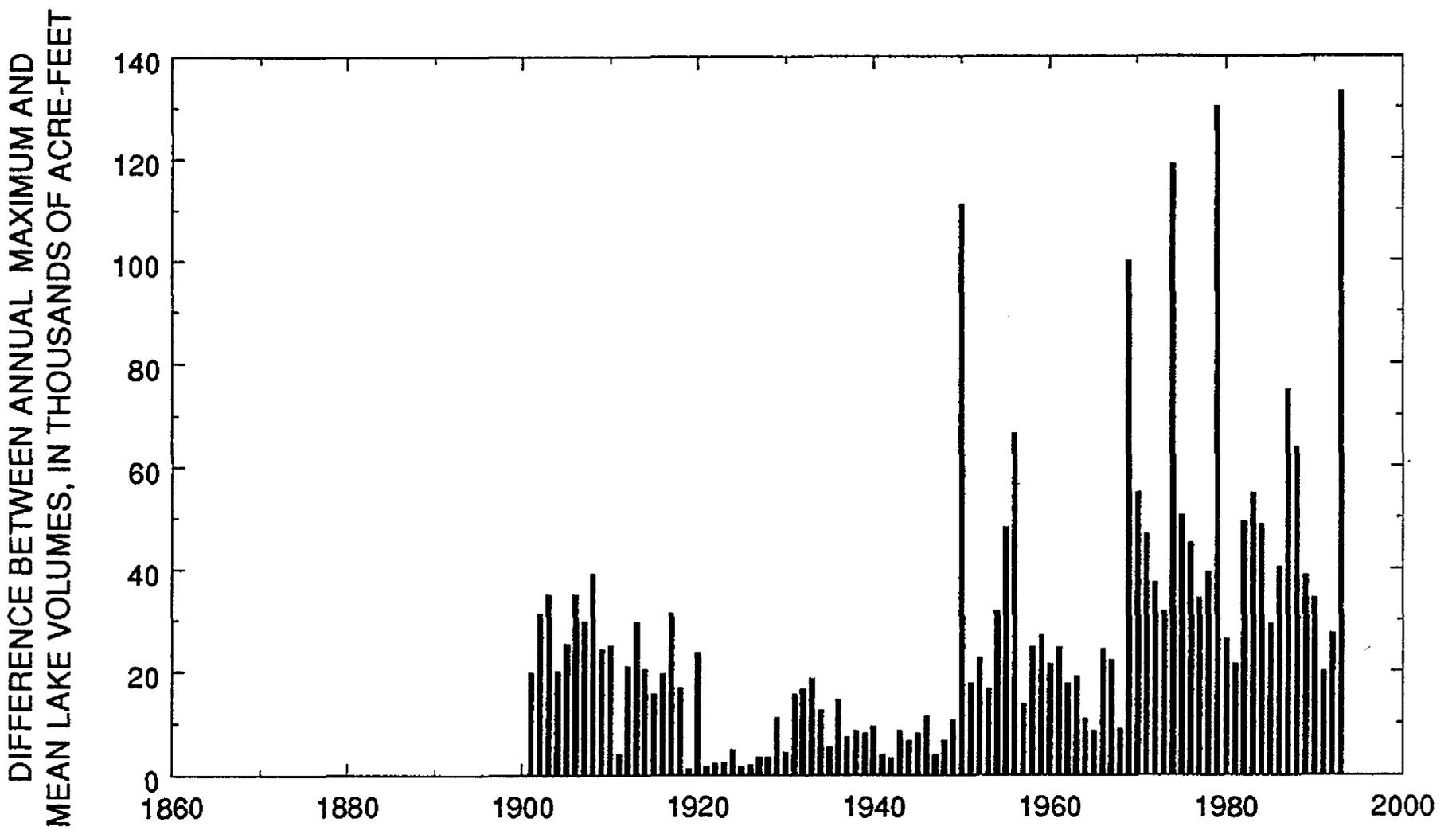


Figure 4. Difference between annual maximum lake volume for year t for Devils Lake and annual mean lake volume for year t , 1901-93.

Data were transformed by use of equation 1 because normality transformations were effective in converting $X(t)$ and $D(t)$ to an approximate Gaussian time series but were not as effective in converting $V_{max}(t)$. Because $D(t)$ is one-half the volume change for year t , it should be highly correlated with inflows during years when inflows are large enough to dominate net evaporation loss.

Annual Lake-Volume Model

Autoregressive moving average (ARMA) models (Box and Jenkins, 1976; Brockwell and Davis, 1987) were adequate for modeling both $X(t)$ and $D(t)$. In order to use standard maximum-likelihood estimation methods for fitting Gaussian ARMA models, power transformations were used to make the residuals from the ARMA models for both $X(t)$ and $D(t)$ as close to normally distributed as possible. The power transformation for $X(t)$ is given by Box and Cox (1964) as

$$X'(t) = \begin{cases} \frac{[X(t)^\lambda - 1]}{\lambda}, \lambda \neq 0 \\ \ln [X(t)], \lambda = 0 \end{cases}, \quad (4)$$

where

$X'(t)$ is the Box-Cox transformed annual mean lake volume for year t , in thousands of acre-feet;
and

λ is the power-transformation parameter.

The Gaussian ARMA model for the transformed series is given by Brockwell and Davis (1987) as

$$X'(t) - \mu = \sum_{j=1}^p \phi_j [X'(t-j) - \mu] + \sum_{j=1}^q \theta_j W(t-j) + W(t), \quad (5)$$

where

μ is the expected value of $X'(t)$;

p is the autoregressive order;

ϕ_j is the j th autoregressive parameter;

q is the moving-average order;

θ_j is the j th moving-average parameter; and

$W(t)$ is a sequence of independent, normally distributed, random variables that have an expected value of zero and a variance of σ_W^2 .

For given values of the autoregressive order, p , the moving-average order, q , and the power-transformation parameter, λ , the statistical package S-Plus (StatSci, 1993) was used to obtain maximum-likelihood estimates of the model parameters (eq. 5) on the basis of recorded values of $X(t)$. Missing values near the beginning of the record (fig. 3) are determined using the ARMA maximum-likelihood

estimation procedure in S-Plus by use of the Kalman filter. The best model (the best values for p , q , and λ) was selected by using the modified Akaike Information Criterion (AIC) statistic,

$$AIC(p, q, \lambda) = -2 \ln(\hat{L}) + 2(p + q + 2) - 2(\lambda - 1) \sum_t \ln[X(t)], \quad (6)$$

where

\hat{L} is the maximum of the likelihood function of the ARMA model fitted to $X'(t)$.

The summation is over all recorded values of $X(t)$. The sum of the first two terms on the right-hand side of equation 6 is the classical AIC statistic (Brockwell and Davis, 1987) and the last term is because of the power transformation (eq. 4). The AIC statistic was computed for all combinations of p , q , and λ values selected from the sets $p \in \{0, 1, 2\}$, $q \in \{0, 1, 2\}$, and $\lambda \in \{1, 1/2, 1/3, 1/4, 1/5, 0\}$. The model that had the lowest AIC statistic had $p = 1$, $q = 1$, and $\lambda = 1/2$. The resulting model is

$$X'(t) - 41.12 = 0.967[X'(t-1) - 41.12] + 0.610W(t-1) + W(t), \quad (7)$$

where

$$X'(t) \text{ is } \frac{X(t)^{1/2} - 1}{1/2}$$

and

$$\hat{\sigma}_W^2 \text{ is } 4.624.$$

The standardized residuals from the fitted ARMA model and a normal probability plot are shown in figure 5. The standardized residuals are defined as $W^*(t) = W(t) / \hat{\sigma}_W$. Data indicate that the square-root transformation is effective in normalizing the model residuals. Graphical analysis of the residual autocorrelation function and the residual periodogram also indicated that the residuals are uncorrelated with zero mean and constant variance (white noise). Standard diagnostics tests, such as the portmanteau test, turning-points test, difference-sign test, and rank test (Kendall's τ), were conducted as described in Brockwell and Davis (1987, chap. 9) to test for nonrandomness in the residuals. The tests indicated no significant autocorrelation, trends, or other deviations from white noise at the 10-percent significance level.

The power transformation for $D(t)$ is given by Box and Cox (1964) as

$$D'(t) = \begin{cases} \frac{[D(t)^\lambda - 1]}{\lambda}, \lambda \neq 0 \\ \ln[D(t)], \lambda = 0 \end{cases} \quad (8)$$

where

$D'(t)$ is the Box-Cox transformed difference between the maximum lake volume for year t and the mean lake volume for year t , in thousands of acre-feet.

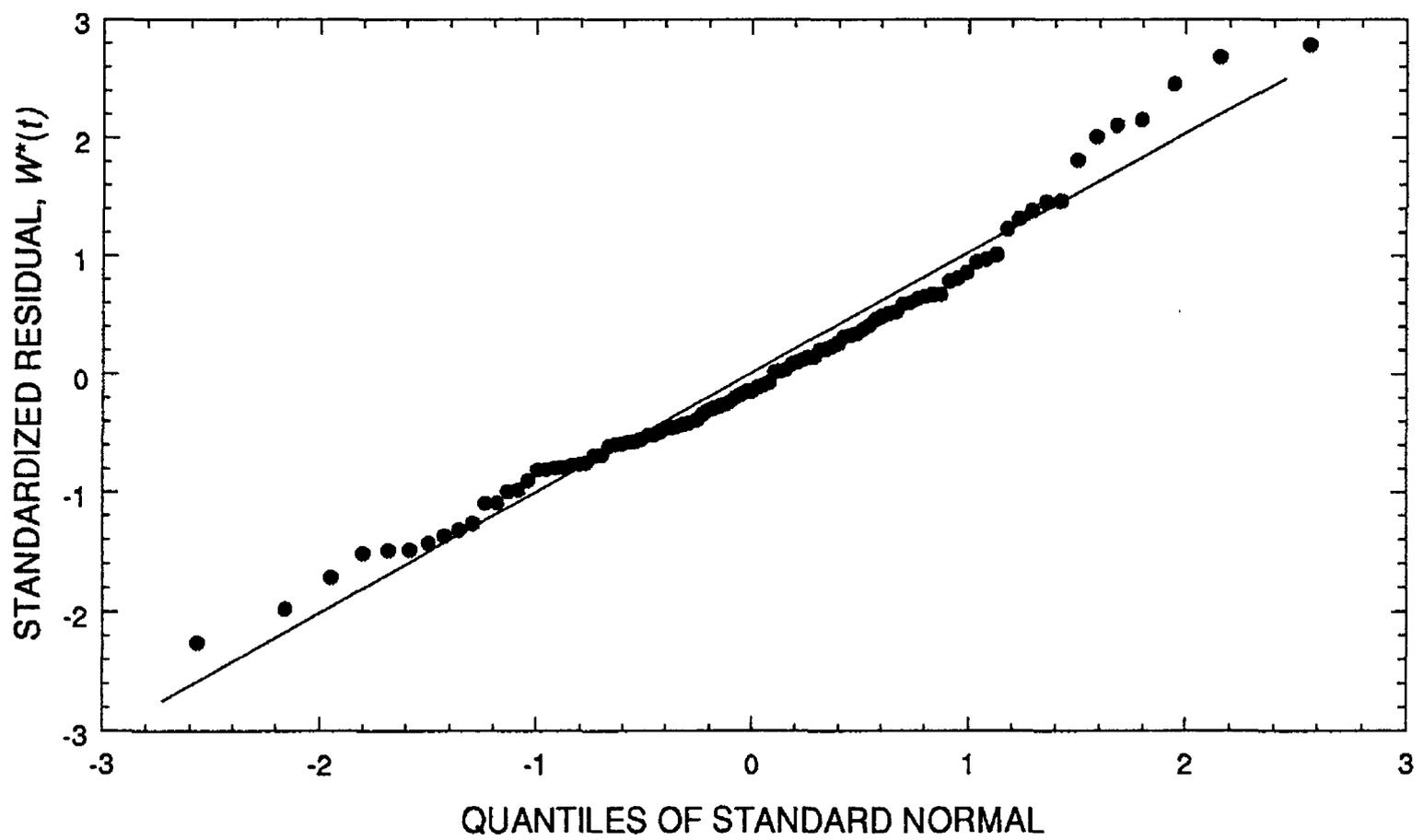
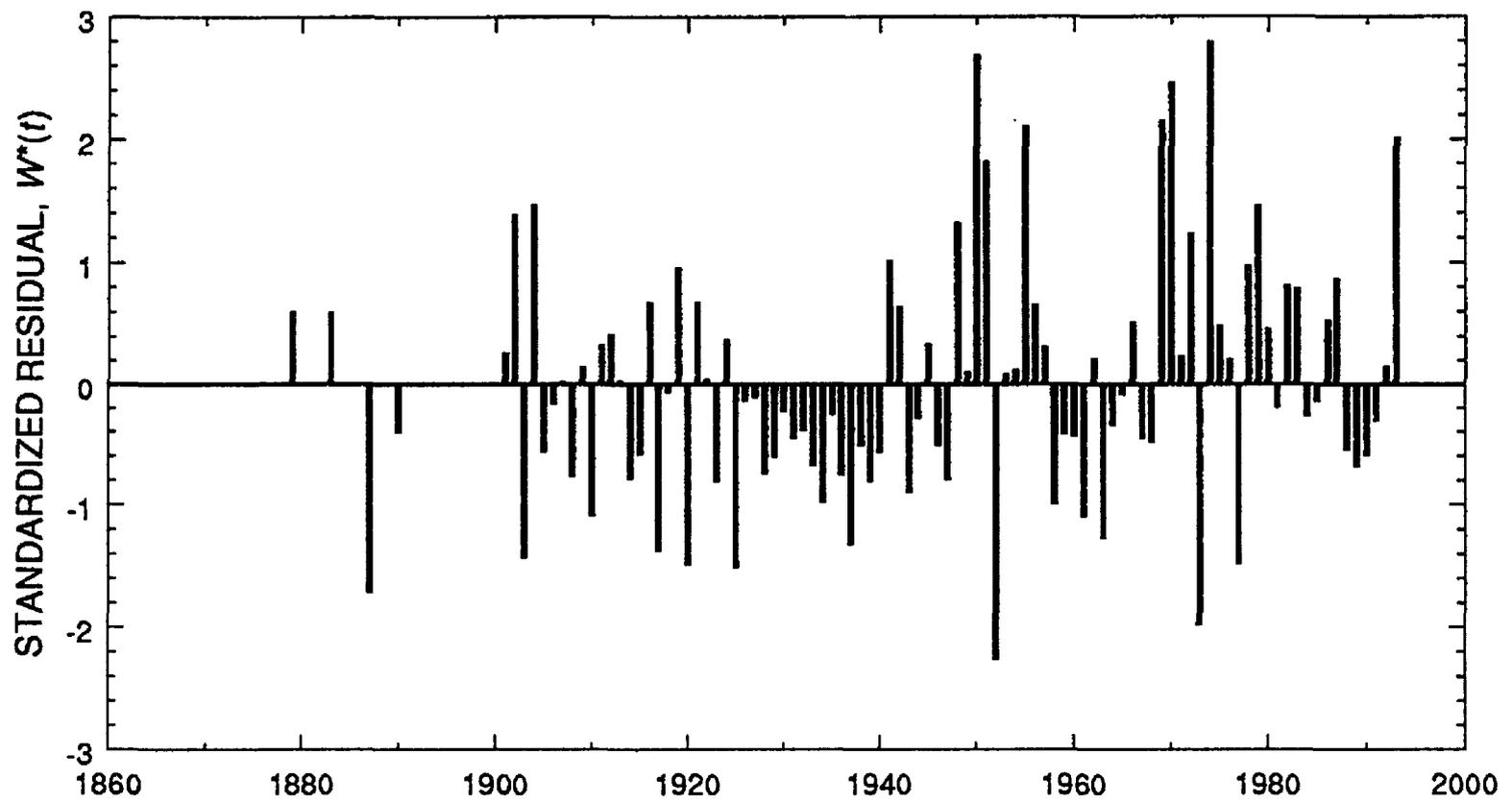


Figure 5. Standardized residuals, $W^*(t)$, from autoregressive moving average model of annual mean lake volumes for Devils Lake, 1867-1993, and normal probability plot.

A Gaussian ARMA model was fitted to various power transformations of $D(t)$ using the same procedure that was used to fit the model to various power transformations of $X(t)$. The ARMA model that had the lowest AIC statistic had $p = 1$, $q = 1$, and $\lambda = 1/4$. The resulting model is

$$D'(t) - 4.912 = 0.943 [D'(t-1) - 4.912] - 0.600Z(t-1) + Z(t), \quad (9)$$

where

$$D'(t) \text{ is } \frac{D(t)^{1/4} - 1}{1/4};$$

$Z(t)$ is a sequence of independent, normally distributed, random variables that have an expected value of zero and a variance of $\hat{\sigma}_Z^2$; and

$\hat{\sigma}_Z^2$ is 2.110.

The standardized residuals from the fitted ARMA model and a normal probability plot are shown in figure 6. The standardized residuals are defined as $Z^*(t) = Z(t)/\hat{\sigma}_Z$. The normality assumption is not satisfied closely in the upper tail of the distribution. Standard diagnostics tests mentioned earlier indicated no autocorrelation, trends, or other nonrandomness in the residuals for $D(t)$. The tests indicated that the residuals are approximately white noise but are not normally distributed. To determine the cause of the nonnormality in the residuals from the ARMA model for $D'(t)$, the relation between the residuals, $W(t)$ and $Z(t)$, from the two models was examined. If the two time series, $X(t)$ and $D(t)$, are statistically independent of one another, then the two residual time series, $W(t)$ and $Z(t)$, should be statistically independent as well. Crosscorrelations between the two residual time series indicated a significant lag-zero crosscorrelation of 0.45. Closer graphical analysis of the residuals indicated that a nonlinear relation exists between $W(t)$ and $Z(t)$. The standardized residuals, $W^*(t)$, for equation 7 were plotted with the concurrent standardized residuals, $Z^*(t)$, for equation 9 (fig. 7). To better understand the nonlinear relation, the interpretation of the model residuals must be understood. Brockwell and Davis (1987) showed that

$$W(t) = X'(t) - E[X'(t)|X'(t-1), X'(t-2), \dots], \quad (10)$$

where

$$E[X'(t)|X'(t-1), X'(t-2), \dots]$$

is the conditional expectation of $X'(t)$ given $X'(t-1), X'(t-2), \dots$;

and

$$Z(t) = D'(t) - E[D'(t)|D'(t-1), D'(t-2), \dots], \quad (11)$$

where

$$E[D'(t)|D'(t-1), D'(t-2), \dots]$$

is the conditional expectation of $D'(t)$ given $D'(t-1), D'(t-2), \dots$

Therefore, in years when $X'(t)$ is much larger than expected on the basis of previous lake levels [i.e., when $W(t)$ is large], $D'(t)$ also tends to be much larger than expected. These years correspond to years when inflows are large in relation to net evaporation. The 5 years that had the largest annual inflows on record (1950, 1969, 1974, 1979, and 1993) are shown in figure 7.

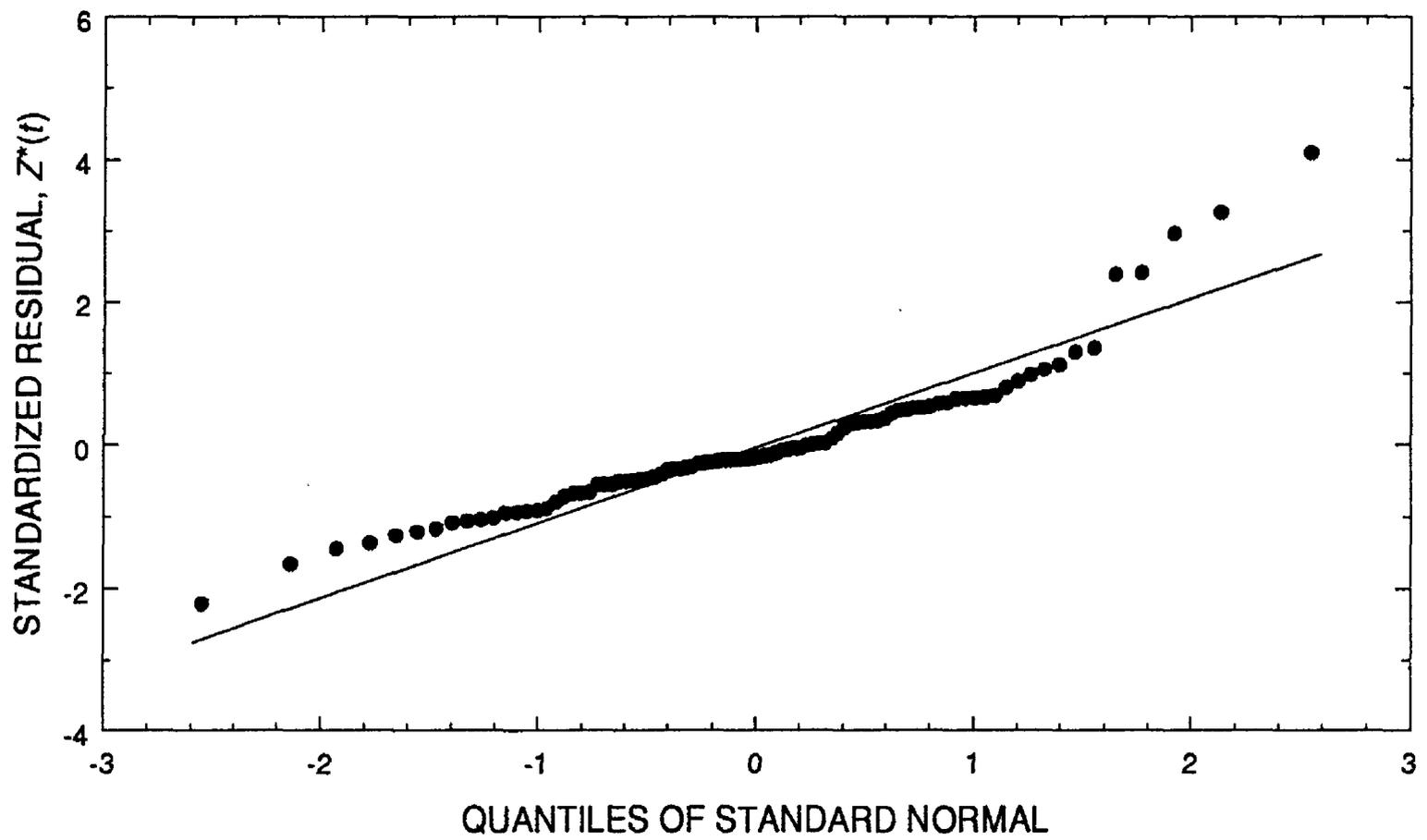
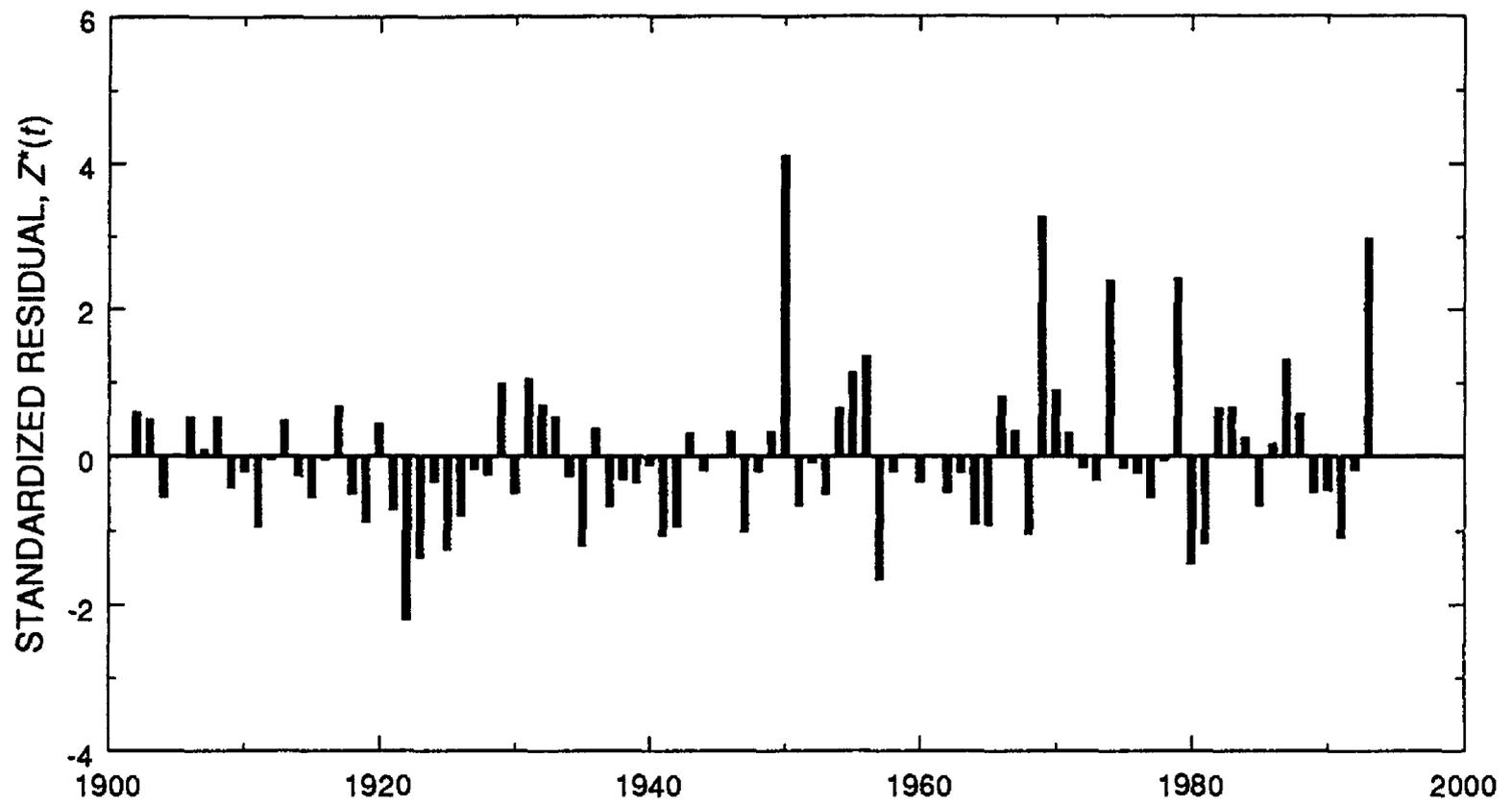


Figure 6. Standardized residuals, $Z^*(t)$, from autoregressive moving average model of differenced annual lake volumes for Devils Lake, 1901-93, and normal probability plot.

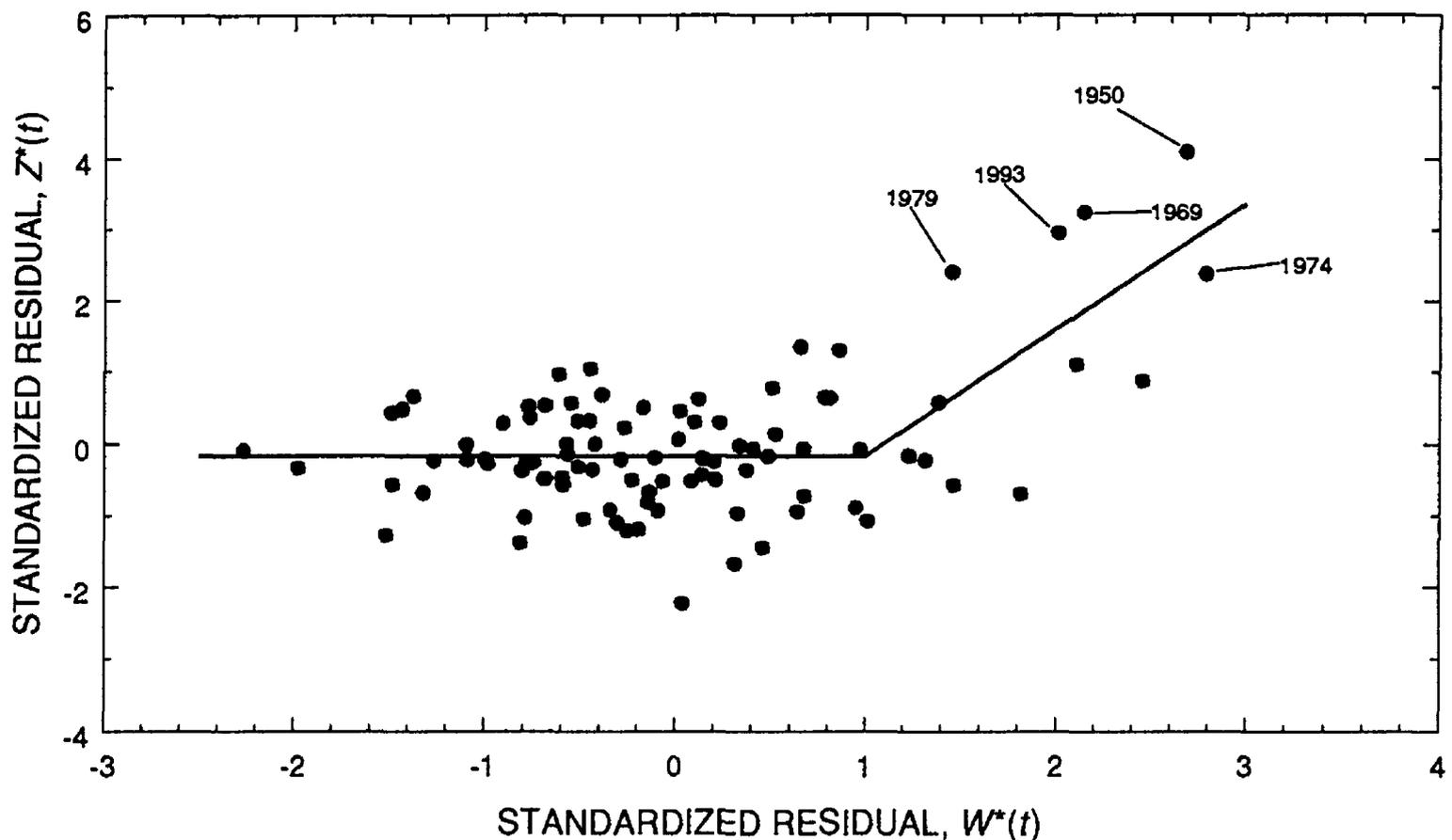


Figure 7. Relation of standardized residuals, $W^*(t)$, from autoregressive moving average model of annual mean lake volumes for Devils Lake to standardized residuals, $Z^*(t)$, from autoregressive moving average model of differenced annual lake volumes.

In order to better model the nonnormality in the residuals for $D'(t)$, the following model was used:

$$Z^*(t) = a_0 + a_1 \max [W^*(t) - 1, 0] + a_2 R^*(t) \quad (12)$$

where

a_0 , a_1 , and a_2 are parameters; and

$R^*(t)$ is a standard normal, random variable that is independent of $W^*(t)$.

Piecewise linear regression (eq. 12) was selected on the basis of figure 7. In order to ensure that $Z^*(t)$ maintains an expected value of zero, which is a very important assumption in the ARMA model (eq. 9), the model was modified to include the following constraint:

$$a_0 = -a_1 E \{ \max [W^*(t) - 1, 0] \} = -0.0881 a_1. \quad (13)$$

The expectation of the truncated standard normal was obtained through numerical integration, and the coefficients in the modified model were estimated by ordinary least-squares regression. The resulting model is

$$Z^*(t) = -0.155 + 1.760 \max [W^*(t) - 1, 0] + 0.779 R^*(t), \quad (14)$$

and the fitted line is shown in figure 7. The adjusted standardized residuals, $R^*(t)$, from equation 14 are shown in figure 8 and are close to a normal distribution. No significant crosscorrelations existed between $W^*(t)$ and $R^*(t)$ out to a maximum lag of 10 years.

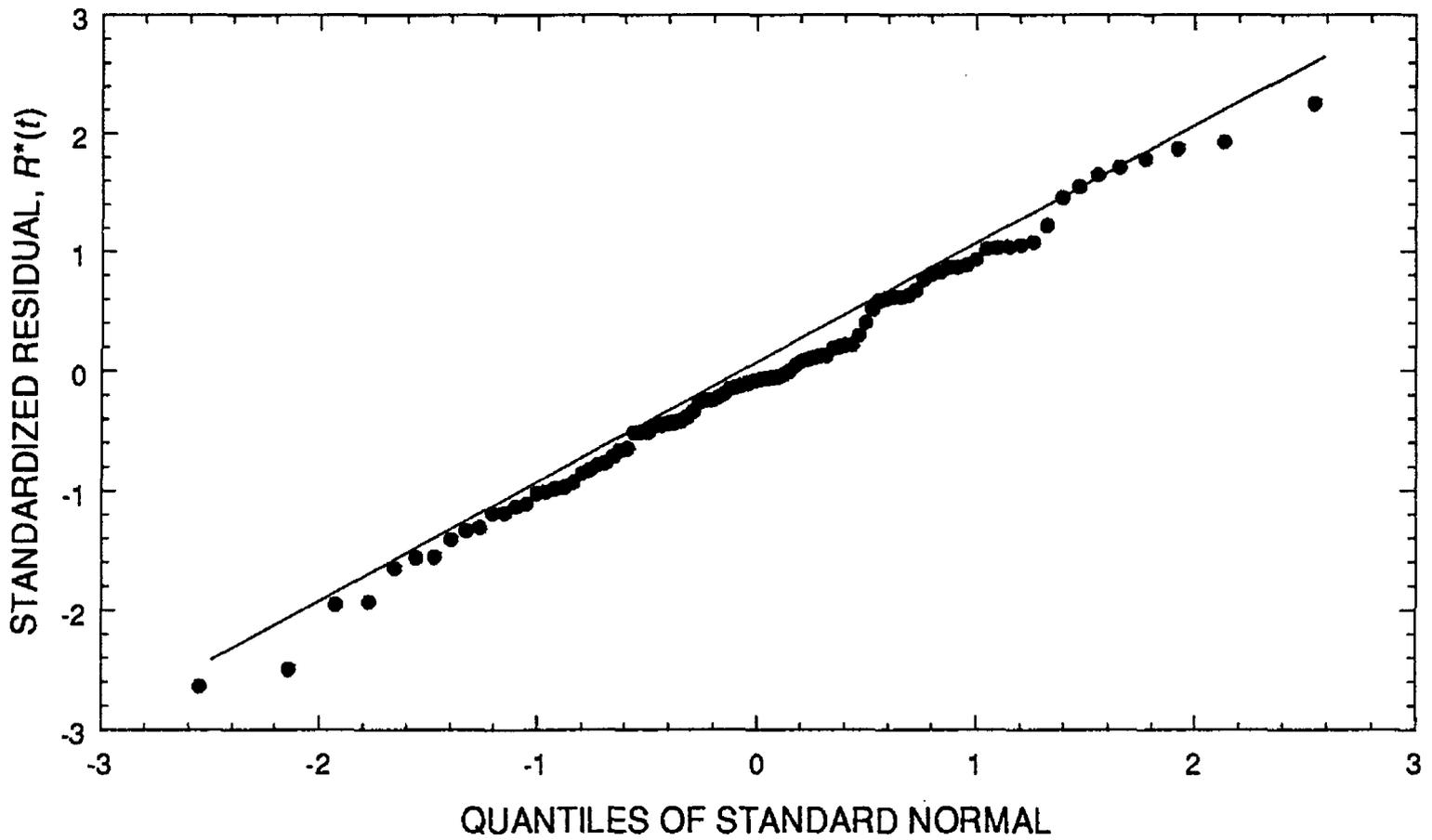
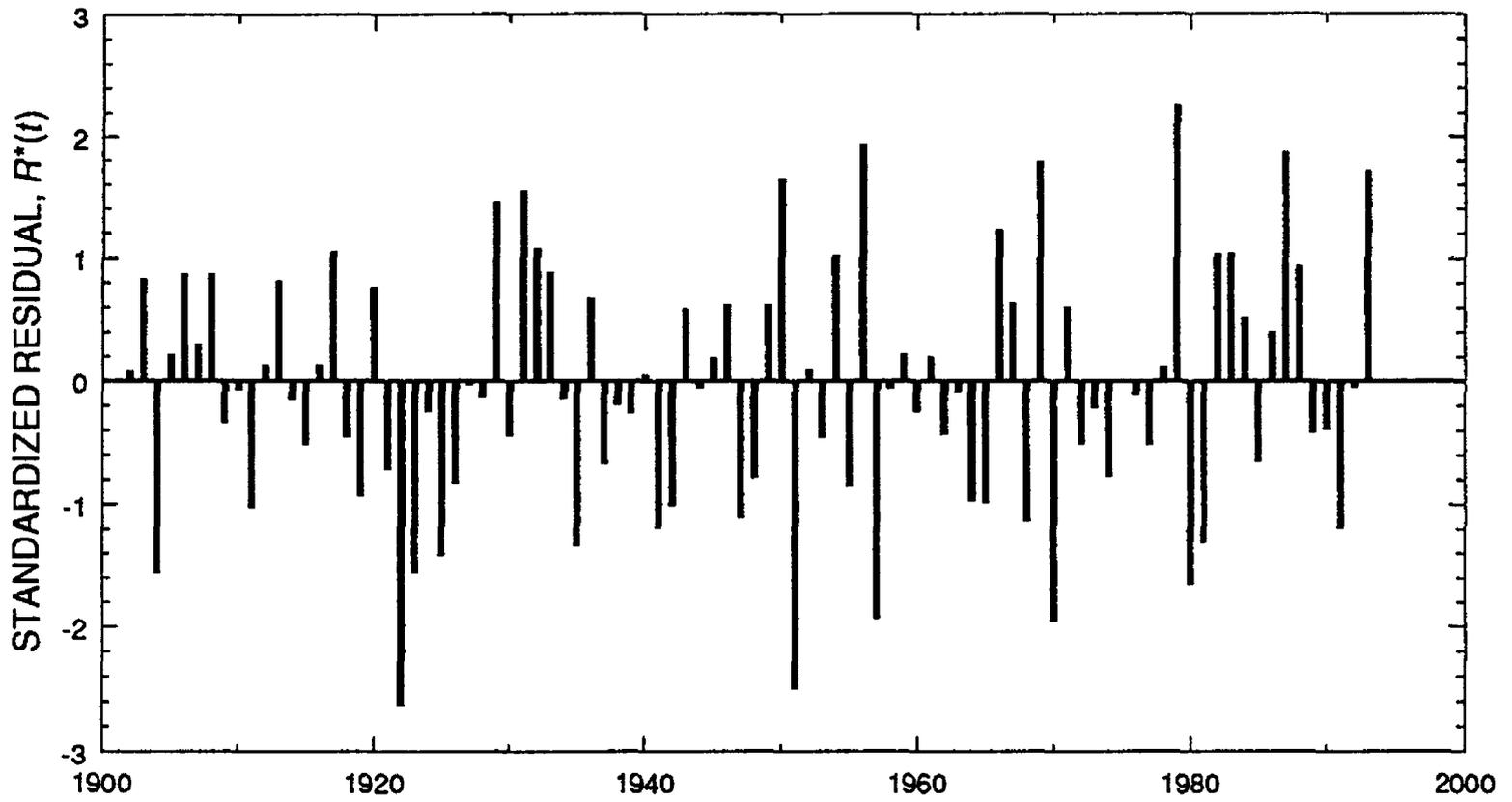


Figure 8. Standardized residuals, $R^*(t)$, from autoregressive moving average model of differenced annual lake volumes for Devils Lake after nonlinear adjustment, 1901-93, and normal probability plot.

Analysis of Model Output

Equation 1, along with the time-series models for $X(t)$ and $D(t)$ [eqs. 7 and 9], can be used to generate future traces of annual maximum lake levels. Before the algorithm for generating lake levels is presented, the difference between conditional and unconditional simulations needs to be outlined. Starting values for X' , W , D' , and Z are required by the ARMA models (eqs. 7 and 9) to generate successive values of X' and D' . Two methods can be used to obtain the required starting values. The first method, called conditional simulation, uses the actual values of the variables from the most recent record (in this study, values are from 1993). $X'(1993)$ and $D'(1993)$ can be obtained directly from the data set, and $Z(1993)$ and $W(1993)$ can be obtained easily from the output of an ARMA model fitting program. In the conditional simulation, all probability statements derived from the simulated data must be interpreted as conditional probabilities given the recorded history of $X(t)$ and $D(t)$ through 1993.

The second method, called unconditional simulation, uses starting values for X' , W , D' , and Z that are selected randomly from their stationary, joint probability distribution. In the unconditional simulation, each trace is assumed to start at a random time. Because of the nonlinear relation (eq. 12) of the model residuals, the exact stationary distribution is difficult to determine analytically. However, approximate unconditional simulations can be obtained by starting each trace in the same way as in the conditional simulation and then dropping the first part of each trace. In the unconditional simulations for this study, the first 200 years of each trace were dropped. Thus, the effects of the conditioning became negligible.

The algorithm for generating conditional simulation lake-level traces is as follows:

1. Generate independent, standard normal, time-series variables, $W^*(t)$ and $R^*(t)$, using a normal, random number generator.
2. Compute the modified standardized residual, $Z^*(t)$, from equation 14.
3. "Unstandardize" the residuals to obtain $W(t)$ and $Z(t)$.
4. Generate the time series of $X'(t)$ from equation 7 and $D'(t)$ from equation 9. Starting values for X' , W , D' , and Z are obtained as described earlier.
5. Reverse the power transformations to obtain $X(t)$ and $D(t)$ in original units.
6. Convert $X(t)$ and $D(t)$ into a time series of annual maximum lake volumes using equation 1.
7. Convert the time series of annual maximum lake volumes to annual maximum lake levels using the Devils Lake elevation-capacity table.

To illustrate the ALV model output, the algorithm was used to generate a single, 1,000-year, unconditional simulation lake-level trace, part of which is shown in figure 9. Because of high temporal persistence in the generated lake-level time series, model performance cannot be measured by how well the model reproduces the statistics of the recorded lake levels of Devils Lake for 1901-93. For example, the 1,000-year, unconditional simulation lake-level trace was used to compute the sample mean of each of 10 nonoverlapping, 100-year lake-level time series. The resulting sample means ranged from 1,410.6 feet to 1,425.4 feet. This indicates extremely high uncertainty in the sample mean of the recorded lake levels, even for a time series of 100 years. Therefore, matching the sample mean lake level for Devils Lake for 1901-93 is not a reliable measure of model performance. The methods used to estimate parameter uncertainty associated with the annual mean lake volume, $X(t)$, and the difference between the annual maximum lake volume and the annual mean lake volume, $D(t)$, are described in supplement 1.

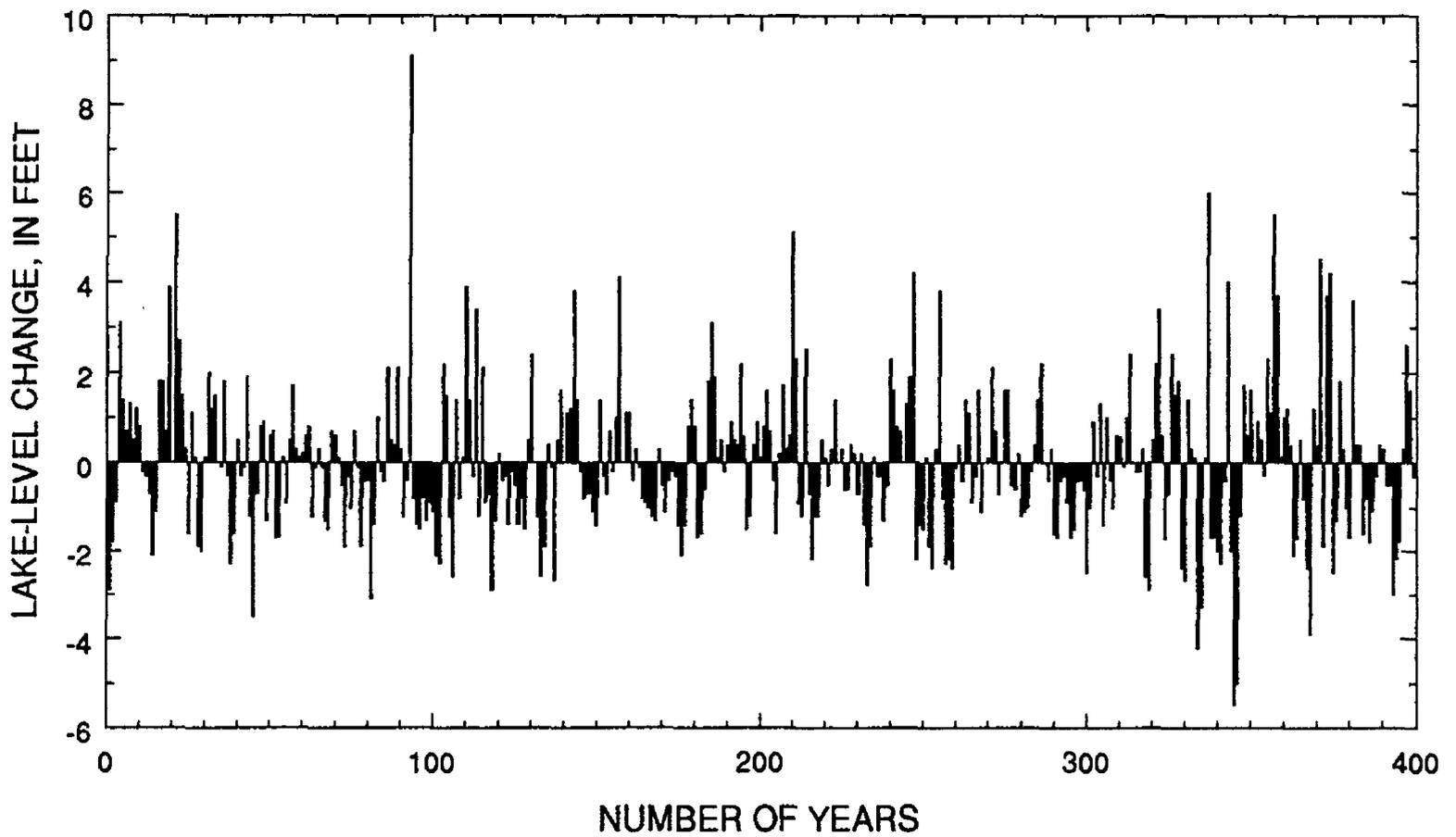
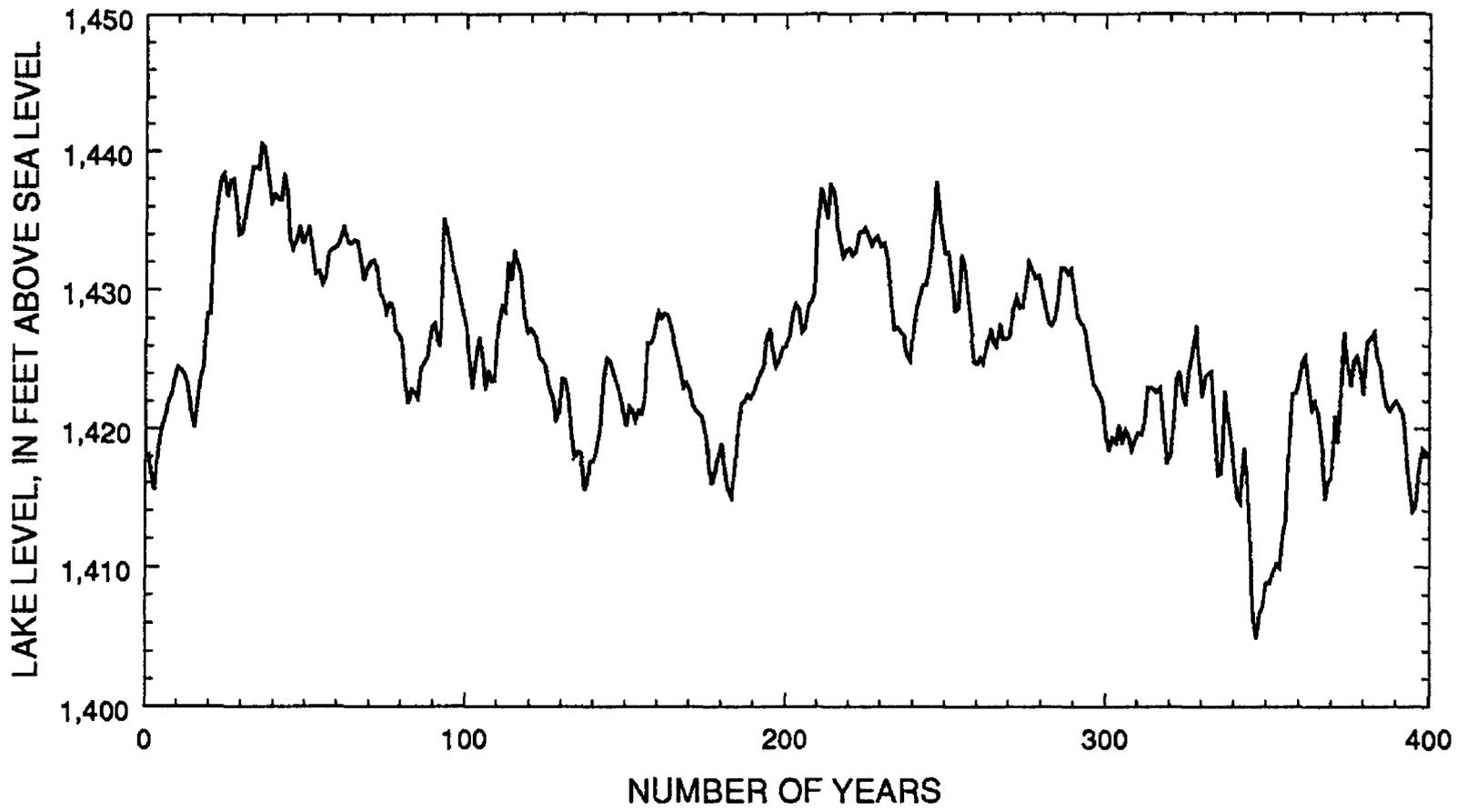


Figure 9. Unconditional simulation lake-level trace from the annual lake-volume model and corresponding lake-level changes for Devils Lake.

Simulated lake-level changes for the ALV model, obtained by taking the difference between the maximum lake level in year t and the maximum lake level in year $t - 1$, also are shown in figure 9. Because the lake-level changes have much less persistence than the lake levels, the model output and the recorded lake levels were compared primarily in terms of lake-level changes. To obtain more information on the simulated lake-level changes and to better compare output from the ALV model and recorded lake-level changes for 1901-93, 100 unconditional simulation lake-level traces, each 93 years in length, were generated from the ALV model. Statistics of the recorded lake-level changes and the simulated lake-level changes are given in table 2. Generally, the model closely reproduces the statistics of the recorded lake-level changes except for the skewness coefficient. The recorded values for 1901-93 are well within the 90-percent range of values generated from the ALV model except for the skewness coefficient. This discrepancy probably was caused by a tendency for the ALV model to produce some unrealistically large lake-level declines. For example, 1 percent of the lake-level changes in the simulated data were less than -3.3 feet (lake-level decline greater than 3.3 feet). The largest recorded lake-level decline for 1901-93 was -2.4 feet. The large lake-level declines in the simulated data occur when the lake is already low (for example, note the large lake-level declines shown between years 300 and 400 in figure 9). The ALV model does not separate inflows to the lake from net evaporation from the lake surface; this can be accomplished only through a detailed WMB model.

Table 2. Selected statistics of recorded lake-level changes for 1901-93 and corresponding values generated from the annual lake-volume model

Statistic	Recorded lake-level change ¹	Annual lake-volume model	
		Average simulated lake-level change ²	90-percent range ³
Mean (feet)	-0.041	-0.016	-0.199 to 0.157
Standard deviation (feet)	1.61	1.58	1.29 to 1.99
Skewness coefficient	1.91	.48	-.13 to 1.22
Lag-1 autocorrelation coefficient	.27	.20	-.08 to .39
Lag-2 autocorrelation coefficient	.01	-.04	-.23 to .10

¹Difference between maximum lake level in year t and maximum lake level in year $t-1$.

²Based on 100 93-year unconditional simulation lake-level traces generated from the annual lake-volume model.

³Numbers represent range that included 90 percent of the values produced from the model.

The algorithm was used next to generate 2,000 conditional simulation lake-level traces, each 50 years in length. These traces were used to demonstrate the effects of conditioning on lake-level frequency analysis. Conditional simulation exceedance levels for Devils Lake for exceedance probabilities of 0.01, 0.25, 0.50, 0.75, and 0.99 are shown in figure 10. The exceedance levels were computed from the 2,000 50-year traces from the ALV model using fixed parameters. Of the 2,000 traces, 1 percent (20 traces) in any given year exceed the upper exceedance level, 25 percent exceed the next level, and so on. The high temporal persistence in the lake levels occurs because the effects of the initial conditions on the subsequent exceedance levels persist for several decades. The unconditional simulation exceedance levels are obtained in the limit as the length of the simulations increases. Because of the high temporal persistence, unconditional simulation frequency analysis is of little use except in dealing with time scales of several centuries.

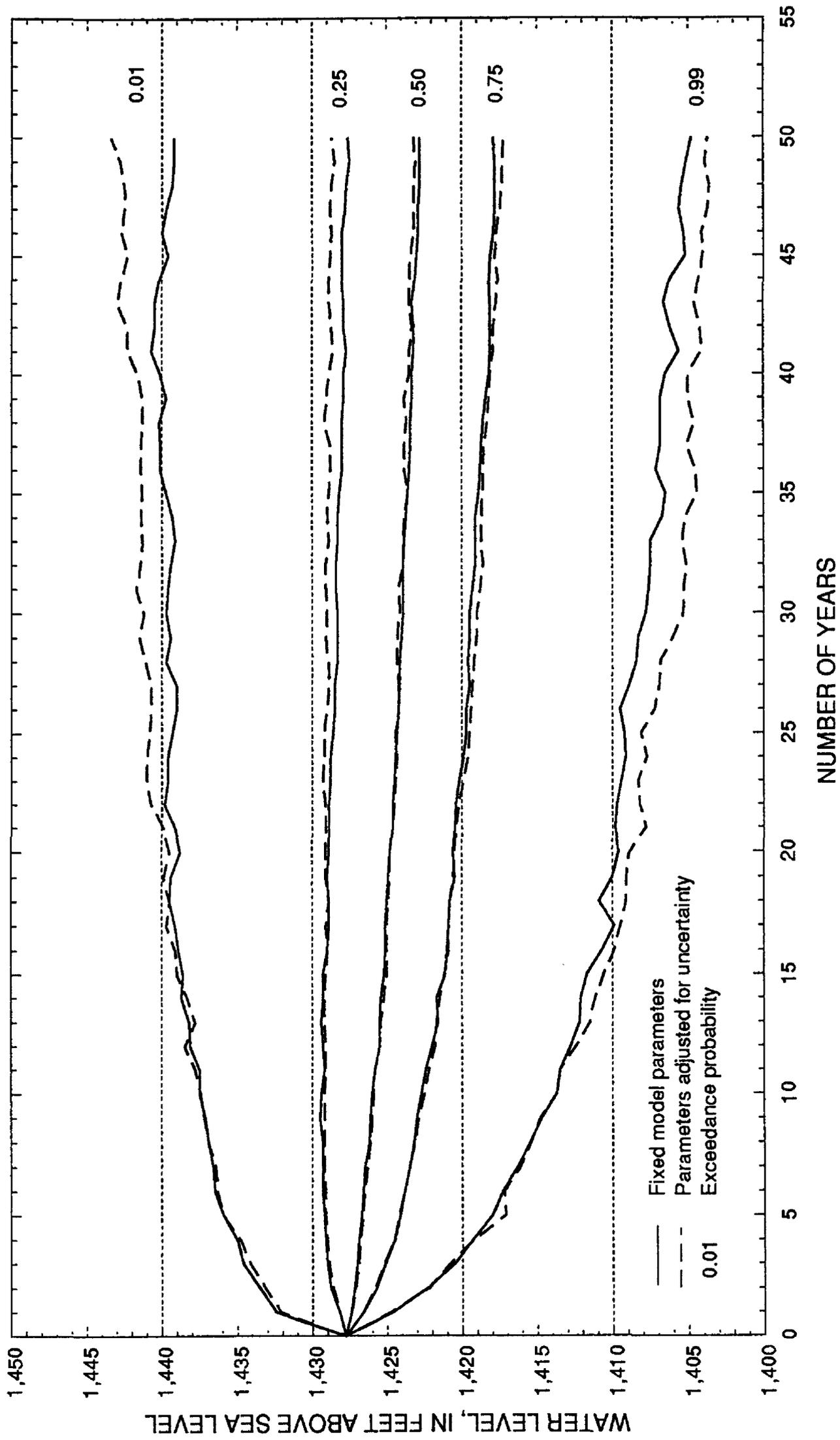


Figure 10. Exceedance levels for Devils Lake computed from the annual lake-volume model using fixed parameters and parameters adjusted for uncertainty. (Year 0 corresponds to 1993.)

Because a finite number of traces (2,000) were used to compute the exceedance levels shown in figure 10, the computed exceedance levels may not reflect the true exceedance levels of the model. Inaccuracies in the computed exceedance levels tend to increase as the exceedance probability becomes closer to zero or 1, as indicated by the roughness in the upper and lower curves in figure 10. To determine accuracies of exceedance levels computed from the 2,000 conditional simulation lake-level traces, nonparametric statistical methods were used to obtain 95-percent confidence intervals for the true exceedance levels (Mood, Graybill, and Boes, 1974, ch. 11). The results obtained for the ALV model using fixed parameters are given in table 3. Values given in the table are the midpoints of a 95-percent confidence interval for the given exceedance level and the half-width of the interval.

Parameter-Uncertainty Analysis

The 2,000 50-year traces (fig. 10) were generated from the ALV model using parameters fixed at their fitted values. This assumes complete confidence in the fitted model. However, because recorded data used to fit the model extend over a short time period in relation to the high temporal persistence of the lake levels, a great deal of uncertainty in the fixed parameters may exist. Stedinger and others (1985) showed that parameter uncertainty can have a large effect on reservoir performance statistics derived from a monthly reservoir simulation model. The approach used in this study is similar to the approach used by Stedinger and others (1985) in that a parameter-uncertainty distribution is developed from which the model parameters for each trace are sampled randomly. Details of the parameter-uncertainty distribution are given in supplement 1. Model parameters included in the uncertainty analysis include those of the ARMA model for $X'(t)$ (eq. 7) and those of the ARMA model for $D'(t)$ (eq. 9). Because no acceptable method exists for including the coefficients from equation 14 in the uncertainty analysis, these coefficients were held fixed. The power-transformation parameters for both $X(t)$ and $D(t)$ also were held fixed.

Conditional simulation exceedance levels for exceedance probabilities of 0.01, 0.25, 0.50, 0.75, and 0.99 are shown in figure 10. The exceedance levels were computed from the 2,000 50-year traces from the ALV model using parameters adjusted for uncertainty. The parameter uncertainty significantly affects the resulting exceedance levels (table 3). The upper exceedance levels are particularly sensitive as indicated by a comparison of the 1-percent exceedance levels (0.01 exceedance probabilities) shown in figure 10. The 1-percent exceedance level from the ALV model that includes parameter uncertainty is 4.6 feet higher in year 50 than the 1-percent exceedance level from the ALV model that has fixed parameters (table 3). The effect becomes even more pronounced at more extreme exceedance levels.

LAKE-LEVEL FREQUENCY ANALYSIS COMPUTED USING A WATER MASS-BALANCE MODEL

Computation of Monthly Precipitation, Evaporation, and Inflow Data

The statistical WMB model requires as inputs seasonal precipitation, evaporation, and inflow data for Devils Lake. Although the final model was fitted on the basis of a quarterly (3-month) time scale, the required inputs were computed using a monthly time scale. A monthly time scale was selected because it is the smallest interval for which reliable estimates of inflow to Devils Lake could be obtained. Monthly inputs were computed for 1950-93 because 1950 is the earliest year in which enough data are available to accurately determine lake evaporation and inflow. Although precipitation data for the city of Devils Lake are available before 1950, these data were not used in this study. The assumption was made that precipitation falling on the lake surface, $P(t)$, during any given month is approximately equal to the

Table 3. Exceedance levels for Devils Lake computed from 2,000 50-year traces from the annual lake-volume model and the water mass-balance model

[Simulation year 1 corresponds to 1994 and exceedance levels are conditioned on recorded data through 1993; number in parentheses is accuracy of the computed exceedance level, in feet; starting lake level for all simulations was 1,427.8 feet above sea level]

Simulation year	Annual lake-volume model		Water mass-balance model		
	Fixed parameters	Parameters adjusted for uncertainty	Fixed parameters	Parameters adjusted for uncertainty	Parameters adjusted for Red River of the North at Grand Forks atreamflow
0.99 exceedance probability					
1	1,424.6 (0.2)	1,424.7 (0.2)	1,425.3 (0.1)	1,427.6 (0.1)	1,427.7 (0.1)
5	1,417.7 (0.8)	1,417.6 (0.8)	1,421.6 (0.3)	1,423.2 (0.4)	1,422.9 (0.4)
10	1,413.6 (1.1)	1,413.3 (1.1)	1,418.5 (0.5)	1,419.3 (0.9)	1,418.7 (0.8)
25	1,408.8 (1.3)	1,407.1 (2.1)	1,414.5 (0.6)	1,414.0 (0.7)	1,412.3 (0.7)
50	1,404.4 (1.6)	1,403.5 (1.1)	1,413.5 (0.7)	1,411.4 (0.9)	1,409.6 (0.6)
0.75 exceedance probability					
1	1,426.6 (0.1)	1,426.6 (0.1)	1,426.7 (0.1)	1,428.2 (0.1)	1,428.1 (0.1)
5	1,424.1 (0.2)	1,424.1 (0.2)	1,425.7 (0.2)	1,426.7 (0.2)	1,426.2 (0.2)
10	1,422.6 (0.2)	1,422.5 (0.3)	1,424.4 (0.2)	1,425.2 (0.2)	1,424.2 (0.2)
25	1,419.7 (0.4)	1,419.5 (0.6)	1,422.6 (0.3)	1,422.7 (0.4)	1,420.7 (0.4)
50	1,417.8 (0.6)	1,417.3 (0.5)	1,421.7 (0.3)	1,421.6 (0.4)	1,419.1 (0.4)
0.50 exceedance probability					
1	1,427.3 (0.1)	1,427.3 (0.1)	1,428.1 (0.1)	1,428.9 (0.1)	1,428.5 (0.1)
5	1,426.5 (0.3)	1,426.4 (0.3)	1,427.8 (0.3)	1,428.7 (0.3)	1,428.2 (0.3)
10	1,425.7 (0.4)	1,425.7 (0.4)	1,427.0 (0.4)	1,428.0 (0.5)	1,426.9 (0.4)
25	1,424.0 (0.6)	1,424.0 (0.7)	1,425.9 (0.5)	1,426.6 (0.5)	1,424.5 (0.5)
50	1,422.4 (0.7)	1,422.8 (0.7)	1,424.9 (0.6)	1,425.7 (0.6)	1,423.0 (0.6)
0.25 exceedance probability					
1	1,428.2 (0.1)	1,428.2 (0.1)	1,429.4 (0.1)	1,430.4 (0.1)	1,430.2 (0.1)
5	1,429.2 (0.2)	1,429.1 (0.2)	1,430.5 (0.3)	1,431.8 (0.3)	1,431.1 (0.3)
10	1,429.4 (0.2)	1,429.2 (0.3)	1,430.4 (0.3)	1,432.0 (0.5)	1,430.7 (0.5)
25	1,428.6 (0.4)	1,429.1 (0.4)	1,429.7 (0.3)	1,431.7 (0.4)	1,429.4 (0.4)
50	1,427.6 (0.4)	1,428.7 (0.4)	1,428.9 (0.3)	1,431.3 (0.5)	1,428.5 (0.5)
0.01 exceedance probability					
1	1,432.3 (0.6)	1,432.3 (0.8)	1,436.4 (1.3)	1,439.2 (1.6)	1,439.2 (1.6)
5	1,436.1 (0.7)	1,435.8 (0.8)	1,440.3 (1.2)	1,444.3 (1.8)	1,443.4 (1.9)
10	1,437.8 (1.0)	1,438.1 (1.3)	1,439.6 (0.9)	1,444.3 (2.1)	1,442.9 (2.2)
25	1,439.6 (0.9)	1,441.0 (0.9)	1,439.7 (1.2)	1,444.4 (1.8)	1,442.5 (2.0)
50	1,439.3 (1.1)	1,443.9 (1.8)	1,440.6 (1.3)	1,445.5 (2.6)	1,443.3 (2.8)

corresponding monthly precipitation total at the city of Devils Lake. Annual precipitation for the city of Devils Lake (U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1951-94) for 1950-93 is plotted in figure 11.

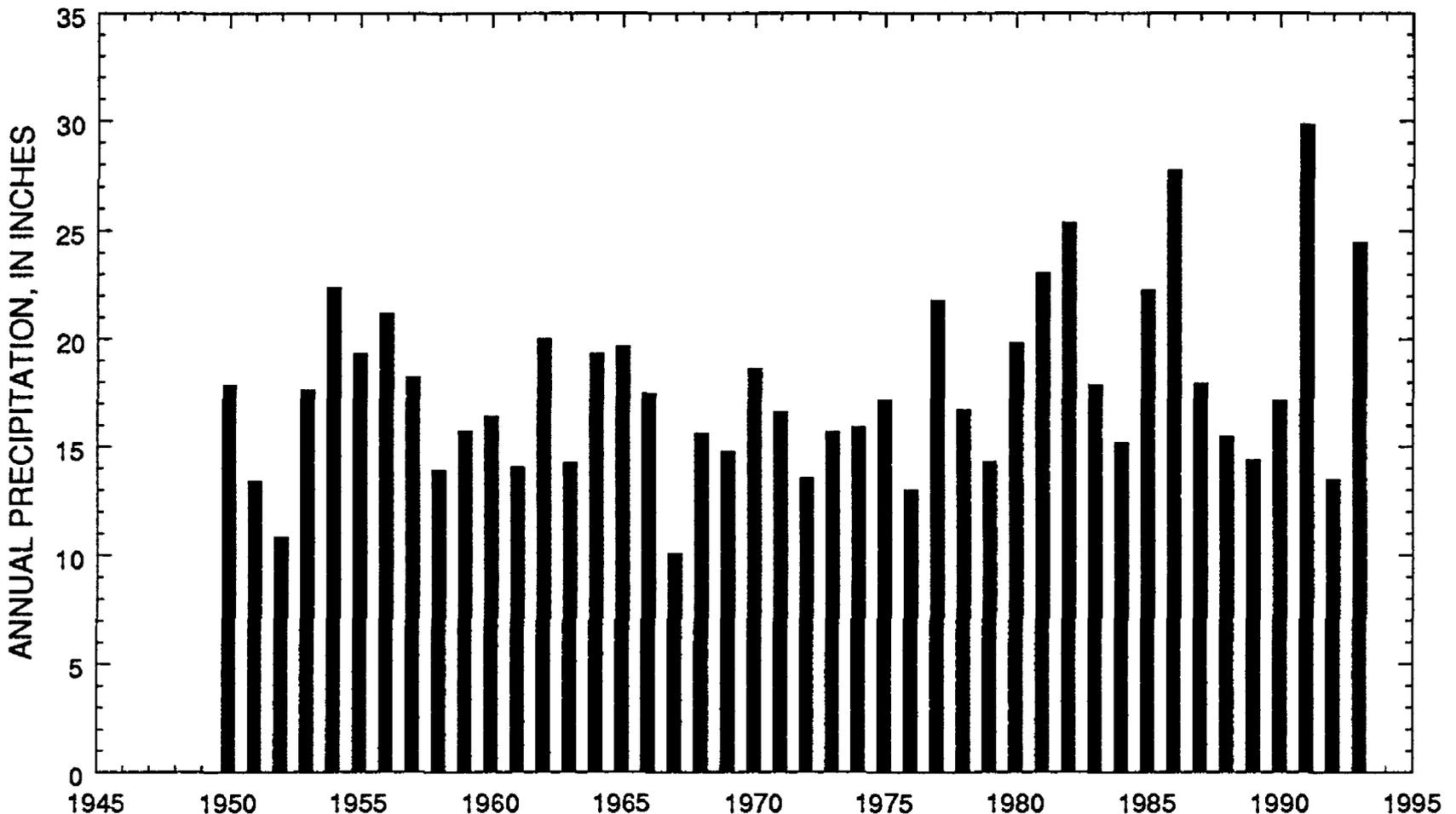


Figure 11. Annual precipitation for the city of Devils Lake, 1950-93.

Monthly lake-evaporation estimates for 1950-93 were computed using pan-evaporation data for Mandan and Dickinson. Ideally, monthly pan-evaporation values for Devils Lake could be adjusted using coefficients developed to convert pan-evaporation estimates for April through September to lake-evaporation estimates. However, because of the relatively short period of record (1951-70) for Class A pan-evaporation data (U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1951-70) for Devils Lake and the large number of missing values for April and May, the Class A pan-evaporation data for Devils Lake were not used to estimate monthly lake evaporation. Little pan-evaporation data are collected during October through March because minimum daily temperatures often are below freezing.

Lake evaporation for April through September was estimated by computing the ratio of the monthly pan evaporation at Mandan and Dickinson to the mean monthly pan evaporation (1950-93) for each station. The mean of the ratios for Mandan and Dickinson was multiplied by the mean percentage of annual lake evaporation for a given month (U.S. Department of Agriculture, Soil Conservation Service, n.d.; table 4) and the product was multiplied by the mean annual lake evaporation for Devils Lake (table 4). Lake evaporation for October through March was estimated by computing the ratio of the total April through September pan evaporation at Mandan and Dickinson to the mean April through September pan evaporation for each station. The mean of the ratios then was multiplied by the percentage of annual lake evaporation that normally occurs in October through March in North Dakota and the product was multiplied by the mean annual lake evaporation. The assumption was made that if the April through September pan evaporation was different than the mean, the October through March pan evaporation differed by the same percentage.

Table 4. Mean monthly percentage of annual lake evaporation and mean monthly lake evaporation for Devils Lake

Month	Mean monthly percentage of annual lake evaporation ¹ (percent)	Mean monthly lake evaporation for Devils Lake ² (Inches)
January	0.75	0.25
February	.95	.32
March	2.30	.77
April	5.85	1.96
May	10.33	3.46
June	13.57	4.55
July	18.59	6.23
August	20.16	6.75
September	14.95	5.01
October	8.53	2.86
November	3.00	1.00
December	1.00	.34
Total	100.0	33.50

¹U.S. Department of Agriculture, Soil Conservation Service [n.d.].

²Farnsworth and Thompson (1982).

Annual lake-evaporation estimates computed using these methods are shown in figure 12. An apparent step-like increase in lake evaporation occurs in 1959. The increase may have been caused by a change in the type of instrument used to measure pan evaporation, a change in the location of the instrument used, or a change in the method used. Swenson and Colby (1955) indicated that Bureau of Plant Industry evaporation pans were used to collect evaporation data at the Mandan and Dickinson Agricultural Experiment Stations. The pans were 6 feet in diameter and 2 feet deep and were buried to within 4 inches of the top of the pan. The coefficient to convert the sunken pan-evaporation values to lake-evaporation values is about 0.95 (Rowher, 1931), but the coefficient to convert the standard Class A pan-evaporation values to lake-evaporation values is about 0.70. Thus, a change from a sunken pan in 1958 to a Class A pan in 1959 could cause the step-like increase in evaporation. However, according to current and former Department of Agriculture employees at the Mandan and Dickinson Agricultural Experiment Stations and meteorologists at the National Weather Service in Bismarck, no changes occurred in the type of instrument used, the location of the instrument used, or the methods used in the late 1950's. Also, no meteorologic equipment changes are mentioned in the annual reports issued for each experiment station. In contrast to these claims, the monthly climatic summaries for North Dakota indicate a Bureau of Plant Industry evaporation pan was used in the late 1950's to collect evaporation data for Mandan and Dickinson (U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1955-64). The summaries indicate a change from the sunken pan to a Class A pan in May 1961 at the Dickinson Experiment Station and in May 1964 at the Mandan Experiment Station.

Multiple linear regression was used to estimate the magnitude of the change between sunken pan-evaporation data and Class A pan-evaporation data for Mandan and Dickinson (U.S. Department of Agriculture, Weather Bureau, 1932a; U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1951-94). The dependent variable was monthly pan evaporation, and the independent variables were the year and a dummy variable. The dummy variable was one for 1950-58 and zero for 1959-92. Parameter estimates for the dummy variable for Mandan range

from 0.69 inch in April to 1.94 inches in August; estimates for Dickinson range from 1.28 inches in September to 4.39 inches in July. Parameter estimates for a given month were added to the pan-evaporation values, and the procedures previously described were used to compute adjusted estimates of lake evaporation for Devils Lake (fig. 13). The maximum lake-evaporation estimate of 42.3 inches (fig. 13) is in good agreement with the computed energy-budget lake-evaporation values of 40.0 inches in 1988 and 41.5 inches in 1989. The energy-budget values are the largest for the period of study, 1986-93.

Selected statistics for monthly Class A pan-evaporation data are given in table 5. Mean annual Class A pan evaporation for May through September was 37.0 inches at Mandan and 40.1 inches at Dickinson.

The relation between monthly Class A pan-evaporation values for Mandan and Dickinson and monthly Class A pan-evaporation values for Devils Lake is shown in figure 14. A linear regression was used to determine if Class A pan evaporation at Mandan and Dickinson is a good estimate of pan evaporation at Devils Lake. The linear regression equations are

$$\text{PANEVAP}_{\text{Devils Lake}} = 0.69 + \text{PANEVAP}_{\text{Mandan}} (0.68), \quad (15)$$

where

$\text{PANEVAP}_{\text{Devils Lake}}$ is estimated monthly pan evaporation at Devils Lake, in inches; and

$\text{PANEVAP}_{\text{Mandan}}$ is recorded monthly pan evaporation at Mandan, in inches;

and

$$\text{PANEVAP}_{\text{Devils Lake}} = 0.97 + \text{PANEVAP}_{\text{Dickinson}} (0.62), \quad (16)$$

where

$\text{PANEVAP}_{\text{Dickinson}}$ is recorded monthly pan evaporation at Dickinson, in inches.

The coefficient of determination for equation 15 is 0.86, and the coefficient of determination for equation 16 is 0.69. Thus, the method used to estimate lake evaporation for Devils Lake should provide a good estimate for the period for which Class A pan-evaporation data are available.

The general equation used to compute monthly inflow to Devils Lake is

$$Q_m(t) = \Delta S_m(t) + \frac{1}{12} E_m(t) A_m(t) - \frac{1}{12} P_m(t) A_m(t), \quad (17)$$

where

$Q_m(t)$ is total inflow for month t , in acre-feet;

$\Delta S_m(t)$ is storage change for month t , in acre-feet;

$E_m(t)$ is evaporation from the lake surface for month t , in inches;

$A_m(t)$ is lake-surface area for month t , in acres; and

$P_m(t)$ is precipitation falling on the lake surface for month t , in inches.

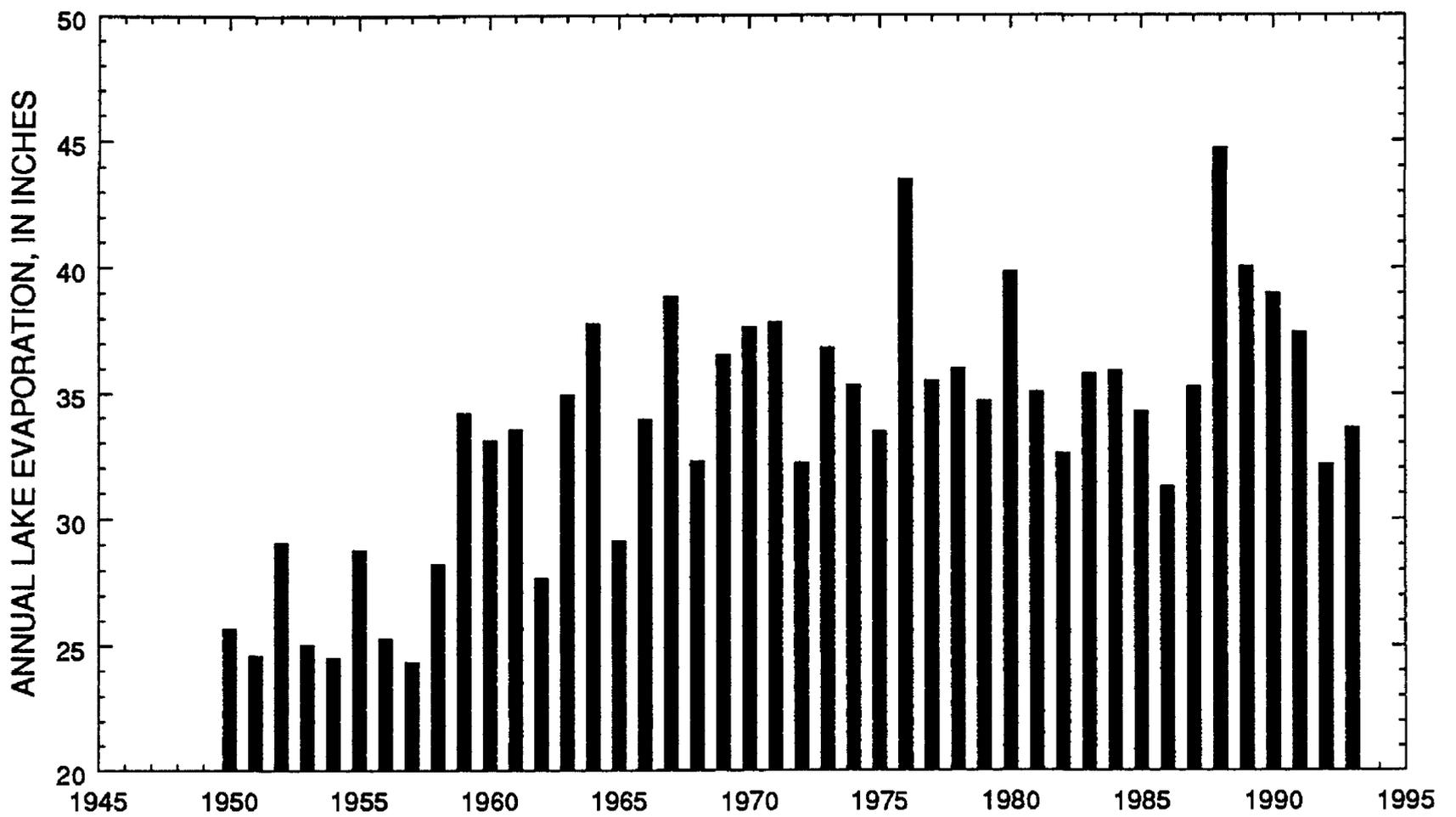


Figure 12. Annual lake evaporation for Devils Lake estimated from unadjusted pan-evaporation data for Mandan and Dickinson, 1950-93.

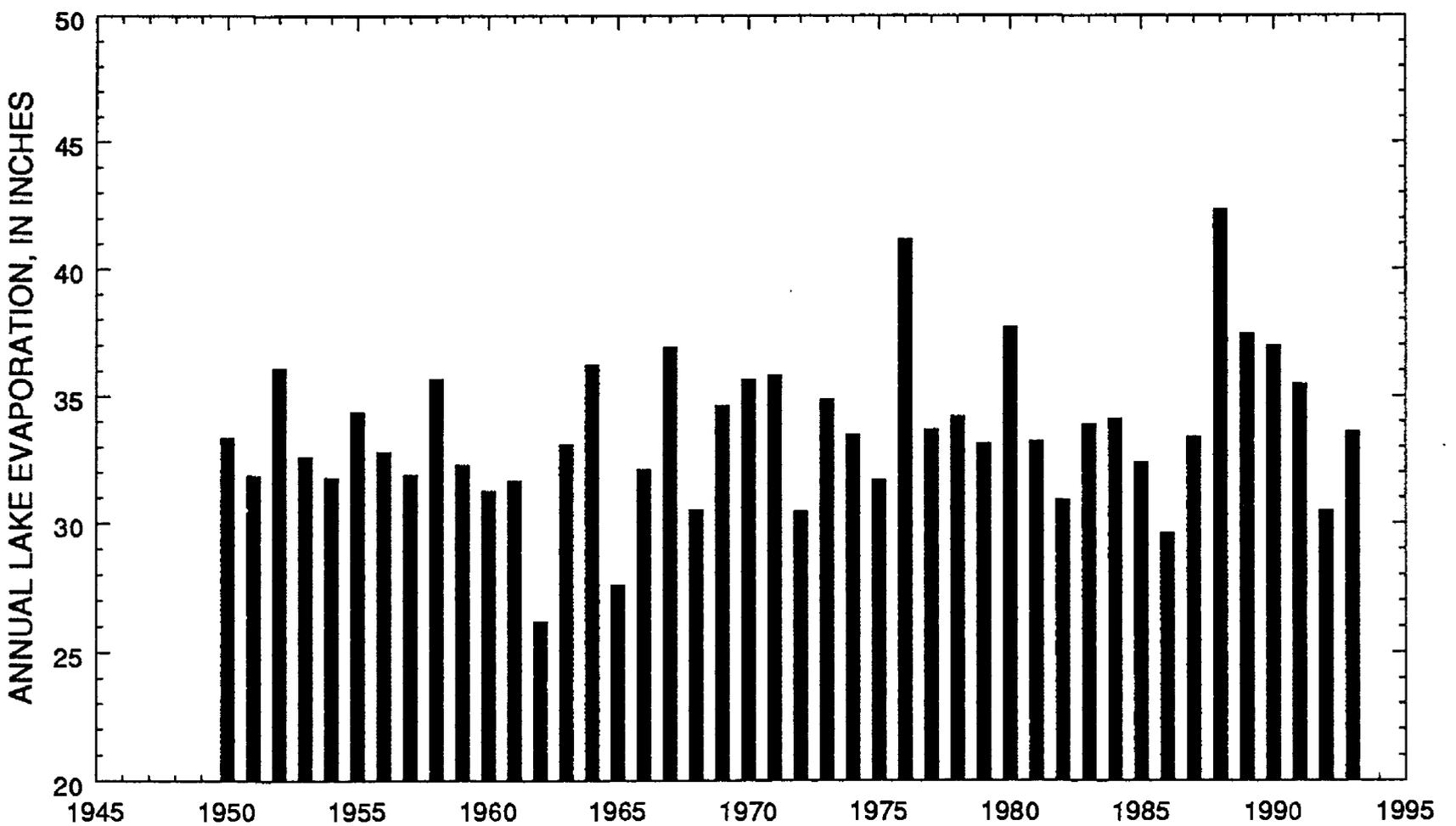


Figure 13. Annual lake evaporation for Devils Lake estimated from adjusted pan-evaporation data for Mandan and Dickinson, 1950-93.

Table 5. Selected statistics for monthly Class A pan-evaporation data for May through September for Mandan and Dickinson

[Data from U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1951-94; evaporation in inches]

Month	Mandan				Dickinson			
	Number of observations	25th percentile	Mean	75th percentile	Number of observations	25th percentile	Mean	75th percentile
May	28	6.20	7.01	7.86	31	5.79	6.96	7.73
June	29	6.50	7.65	8.49	31	7.27	8.35	9.01
July	29	8.37	8.99	9.70	31	8.65	9.87	10.8
August	29	6.95	8.00	8.99	30	8.03	9.17	10.2
September	29	4.65	5.23	5.77	30	5.28	5.72	6.56
Total for May through September	28	34.2	37.0	39.0	30	37.1	40.1	41.8

Annual inflow to Devils Lake for 1950-93 is shown in figure 15. The WMB model (eq. 17) was used to compute the monthly inflow to Devils Lake. The monthly storage change was determined by subtracting the capacity of the lake on the last day of the month from the capacity on the first day of the month. Elevation-area-capacity tables for Devils Lake were obtained from the North Dakota State Water Commission. Monthly estimates of evaporation and precipitation were based on the methods previously discussed.

Small negative inflows [$Q_m(t)$] can occur, especially in the summer when $Q_m(t)$ is near zero. Negative inflows are caused by relatively small errors in the estimates of $\Delta S_m(t)$, $E_m(t)$, and $P_m(t)$. For example, if $Q_m(t)$ and $\Delta S_m(t)$ are zero, $E_m(t)$ is known, $A_m(t)$ is 60,000 acres, and the estimate for $P_m(t)$ is only 0.50 inch larger than the true $P_m(t)$, inflow is -2,500 acre-feet.

When the lake level of Devils Lake is higher than 1,425 feet above sea level, Devils Lake, East Bay of Devils Lake, and East Devils Lake are one continuous water body and one water balance can be used to estimate $Q_m(t)$. However, when the lake level is less than 1,425 feet above sea level, separate water balances are needed to estimate $Q_m(t)$. For example, during 1950-79, separate water balances were computed for Devils Lake east of Highway 20, East Bay of Devils Lake, and East Devils Lake. Only sporadic lake levels are available for East Bay of Devils Lake for the period, and about four lake levels per year are available for East Devils Lake. Thus, linear interpolation was used to determine monthly lake levels needed to compute the ΔS_m and A_m terms in equation 17.

During 1950-79, outflow from Devils Lake to East Bay of Devils Lake occurred whenever the lake level of Devils Lake was greater than about 1,415 feet above sea level. Estimates of outflow from Devils Lake to East Bay during 1950-79 were needed to complete the water balance for East Bay. The estimates were obtained from Swenson and Colby (1955) and the North Dakota State Water Commission (Dale Frink, oral commun., 1994). During 1979, a large volume of snowmelt inflow filled East Bay and East Devils Lake and, by October 1, 1979, the lakes were at the same level.

From October 1979 through July 1989, one water balance was used to estimate the total inflow to Devils Lake. By August 1989, Devils Lake had declined to 1,425 feet above sea level and East Devils

Lake was separated from Devils Lake. The lake level of Devils Lake was less than 1,425 feet above sea level until September 1993. From August 1989 through August 1993, two water balances were used to estimate the total inflow; from September through December 1993, one water balance was used to estimate the total inflow.

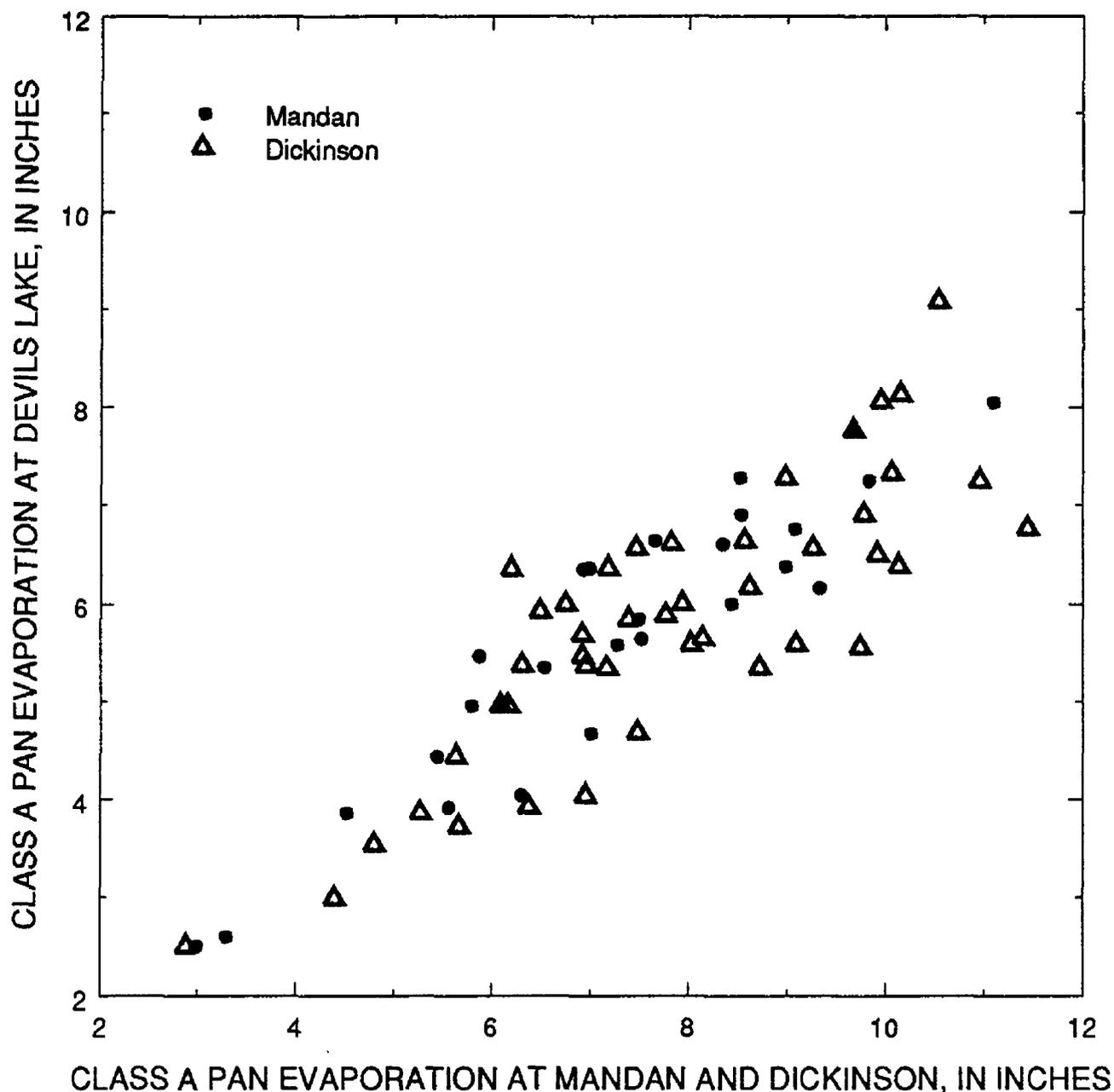


Figure 14. Relation of Class A pan evaporation at Mandan and Dickinson to Class A pan evaporation at Devils Lake, 1951-71.

Statistical Tests for Trends and Long-Range Dependence

The statistical time-series model that will be used to model the recorded precipitation, evaporation, and inflow assumes that no trends or long-range dependence exists in the inputs. In order to test this assumption, the time series of annual precipitation (fig. 11), evaporation (fig. 13), and inflow (fig. 15) were analyzed.

The sample autocorrelation functions for the three annual time series are given in figure 16. Annual inflow was log-transformed because of the large skewness coefficient for untransformed annual inflow (the skewness coefficient for untransformed annual inflow is 2.13; the skewness coefficient for log-transformed inflow is 0.01). The annual precipitation and evaporation time series show no significant autocorrelations out to a maximum lag of 10 years. However, because a significant lag-1 autocorrelation of 0.36 exists in

the log-transformed inflow, the trend tests for inflow are applied to the residuals from a first-order autoregressive model for log-transformed annual inflow. Kendall's *tau*, Spearman's *rho*, and Pearson's *r* were used to test for trends as described in Helsel and Hirsch (1992, ch. 8), where time is used as the independent variable. Test results are given in table 6.

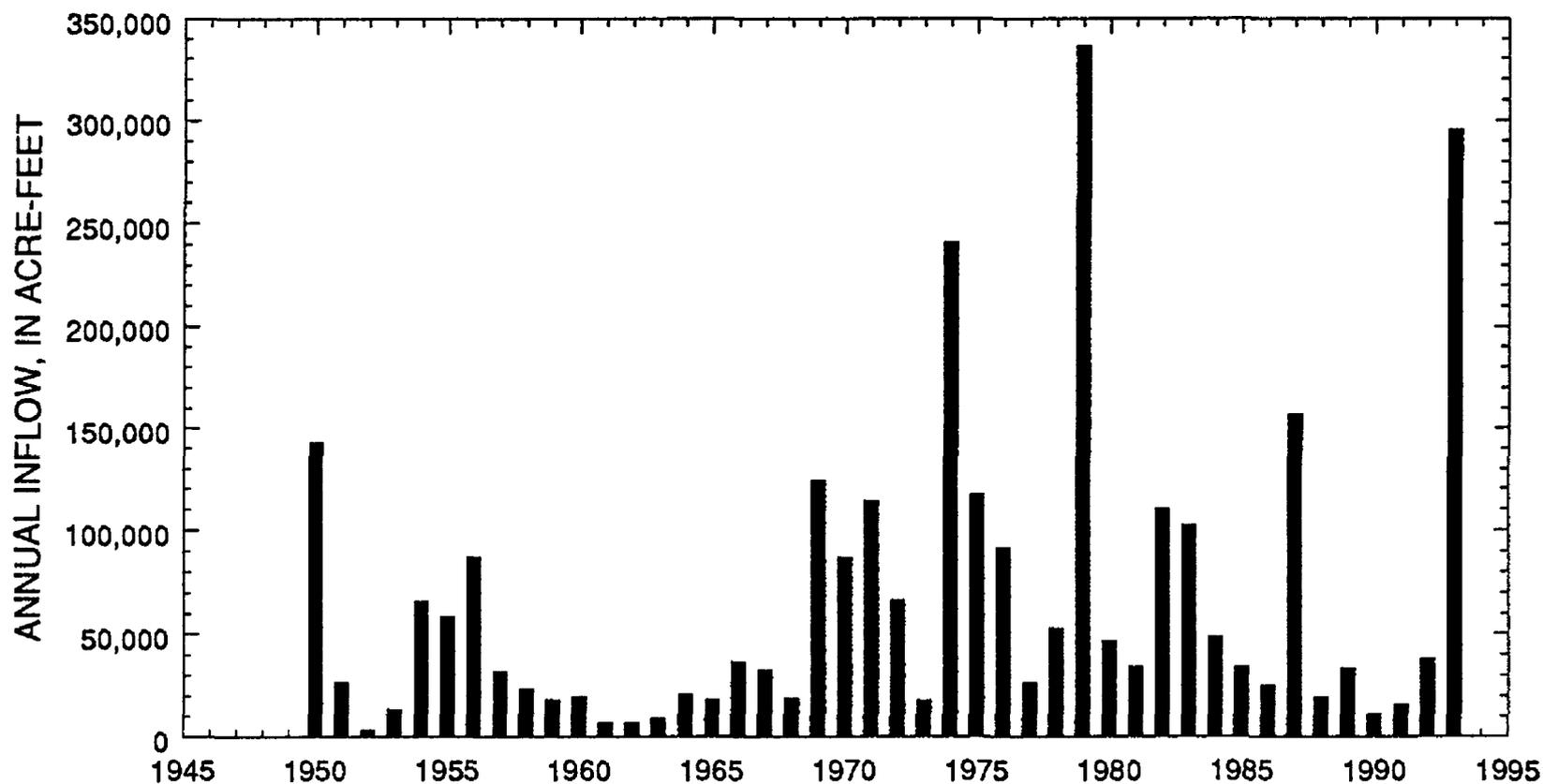


Figure 15. Estimated annual inflow to Devils Lake, 1950-93.

The *p* values for Pearson's *r* for the annual precipitation and inflow time series indicate that *r* is significantly different from zero at the 10-percent level. However, the larger *p* values for Kendall's *tau* and Spearman's *rho* indicate no upward trend in precipitation or inflow. Assuming an upward trend in precipitation or inflow when no such trend exists would have serious consequences on the WMB model. Therefore, the annual precipitation, evaporation, and inflow time series were assumed to be stationary for this study.

The completion of Channel A in 1979 may have significantly altered the distribution of inflow even though the trend tests (table 6) are not sensitive enough to detect such a change. Therefore, the Wilcoxon rank-sum test was used to test for a step trend in 1979. The *p* value of the test is 0.279, indicating that any changes in the distribution of inflow caused by Channel A are not large enough to be detected using standard statistical procedures.

Table 6. Selected statistics of trend tests for annual precipitation, evaporation, and inflow for Devils Lake, 1950-93

Precipitation	Evaporation	Inflow ¹
<i>tau</i> = 0.13, <i>p</i> = 0.21	<i>tau</i> = 0.13, <i>p</i> = 0.22	<i>tau</i> = 0.16, <i>p</i> = 0.13 ¹
<i>rho</i> = 0.12, <i>p</i> = 0.21	<i>rho</i> = 0.22, <i>p</i> = 0.15	<i>rho</i> = 0.23, <i>p</i> = 0.13 ¹
<i>r</i> = 0.30, <i>p</i> = 0.05	<i>r</i> = 0.24, <i>p</i> = 0.12	<i>r</i> = 0.29, <i>p</i> = 0.06 ¹

¹Based on residuals from autoregressive model of log-transformed annual inflow.

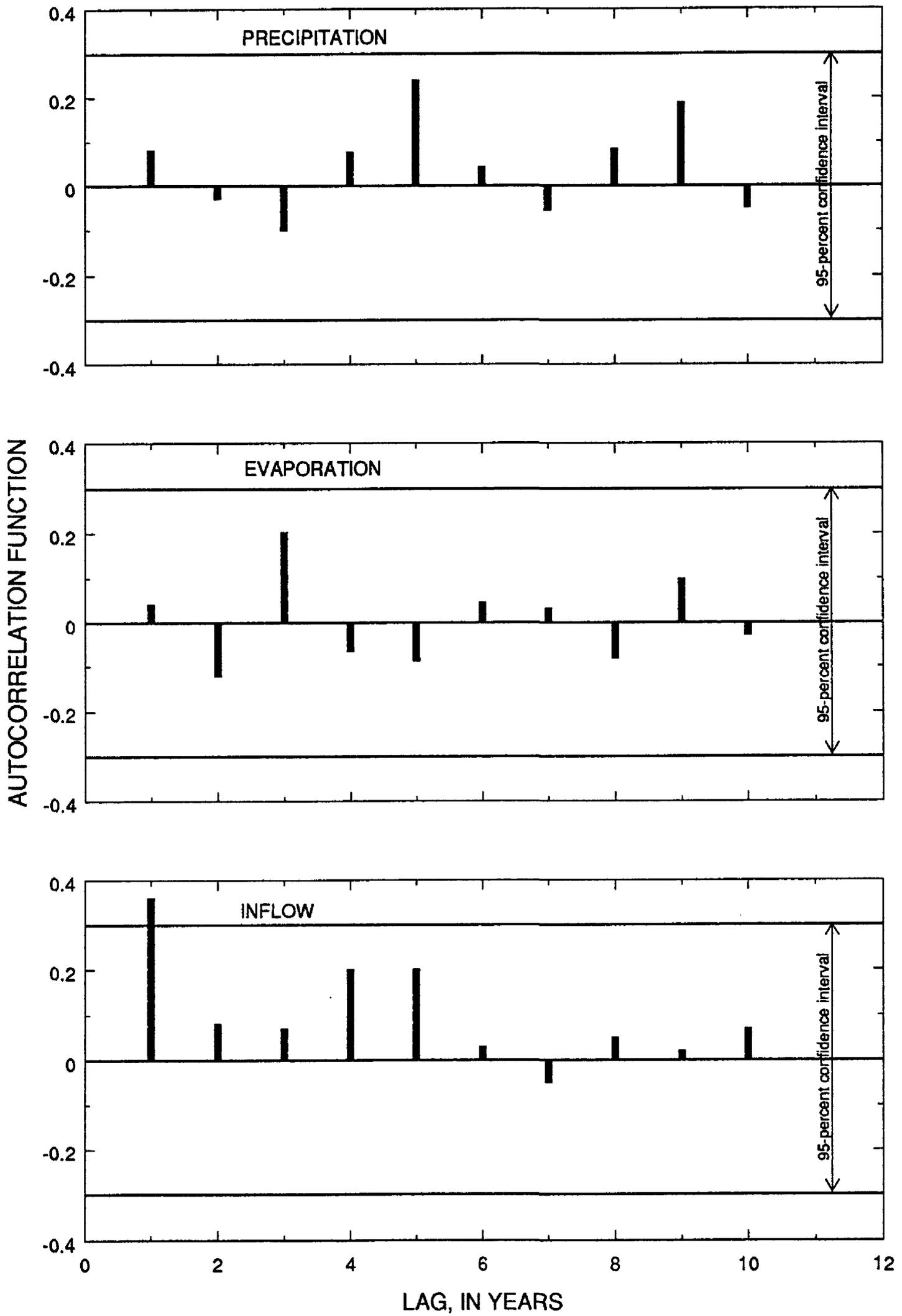


Figure 16. Autocorrelation functions for annual precipitation, annual evaporation, and natural logarithm of annual inflow for Devils Lake.

Long-range dependence, often referred to in hydrologic literature as the Hurst effect (Hurst, 1951), can have a large effect on statistical properties of inputs to the WMB model and, therefore, on resulting lake-level frequency relations (Bowles and James, 1986). Fractional Gaussian noise was proposed by Mandelbrot and Van Ness (1968) as a time-series model for reproducing long-range dependence. Davies and Harte (1987) derived the locally optimal test of the null hypothesis that a given time series is fractional Gaussian noise with Hurst parameter $H = 1/2$ (Gaussian white noise) versus the alternative hypothesis where $H > 1/2$ (long-range dependence). Davies and Harte (1987) showed that the locally optimal test is more efficient for detecting long-range dependence than the widely used rescaled range test proposed by Mandelbrot and Wallis (1969). The Davies and Harte test statistic (1987), R , is given as

$$R = \frac{\sum_{i=1}^n \sum_{j=1}^n [X(i) - \bar{X}] [X(j) - \bar{X}] T(i, j)}{\sum_{i=1}^n [X(i) - \bar{X}]^2}, \quad (18)$$

where

$X(i)$ is a time series of length n ;

\bar{X} is the sample mean; and

$T(i, j)$ is

$$\begin{cases} 0; & i = j \\ \log 2; & i = j \pm 1 \\ \frac{1}{2} \left[|i - j| \log \{1 - (i - j)^{-2}\} + \log \left\{ \frac{(|i - j| + 1)}{(|i - j| - 1)} \right\} \right]; & (|i - j| > 1). \end{cases} \quad (19)$$

Long-range dependence is indicated by a large value of R . The p value of the test is obtained through numerical simulation as described by Davies and Harte (1987). The test statistic, R (eq. 18), and the corresponding p value were computed for the annual precipitation time series ($R = 0.122$ and $p = 0.11$), the annual evaporation time series ($R = 0.104$ and $p = 0.13$), and the first-order autoregressive residuals for log-transformed annual inflow ($R = 0.134$ and $p = 0.11$). No test statistics were significant at the 10-percent level. Thus, for this study, the assumption was made that no long-range dependence exists in the inputs.

Water Mass-Balance Model

A comparison of lake-level traces generated from monthly and quarterly (3-month) WMB models indicated that the lake-level traces generated from the quarterly model were almost identical to the traces generated from the monthly model. Therefore, the WMB model used in this study was based on quarterly rather than monthly data. Quarterly precipitation, evaporation, and inflow time series obtained by aggregating monthly data are shown in figure 17. The quarters will be referred to as winter (January,

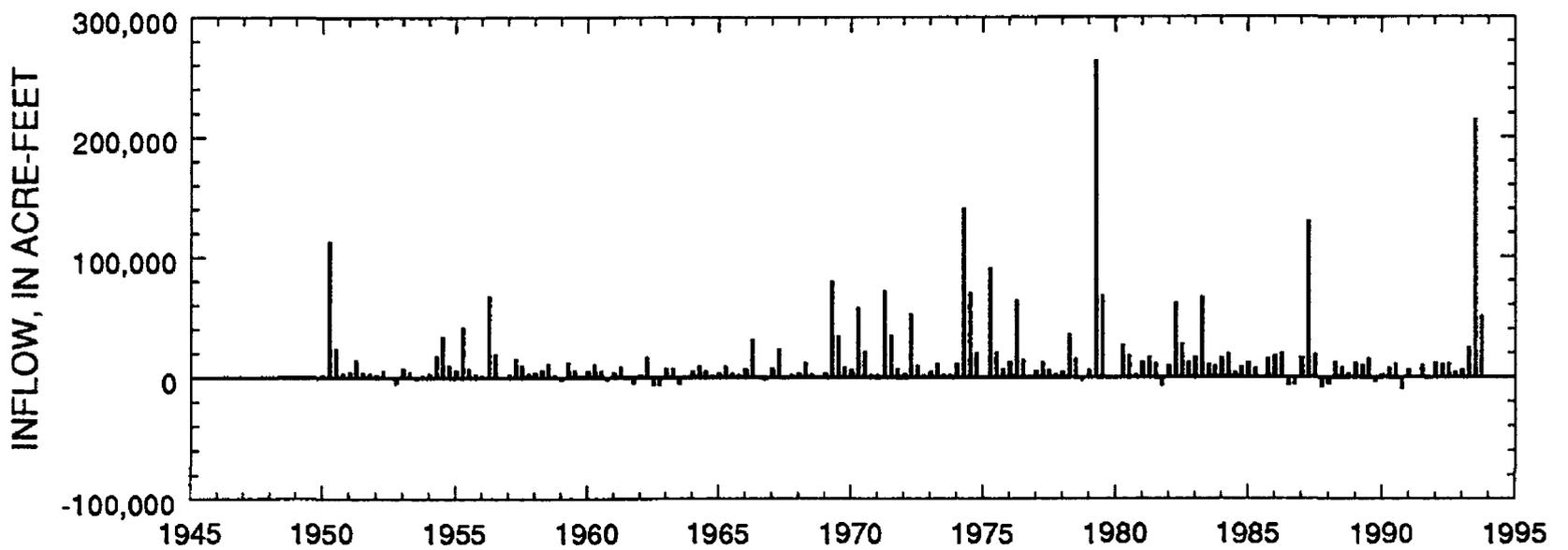
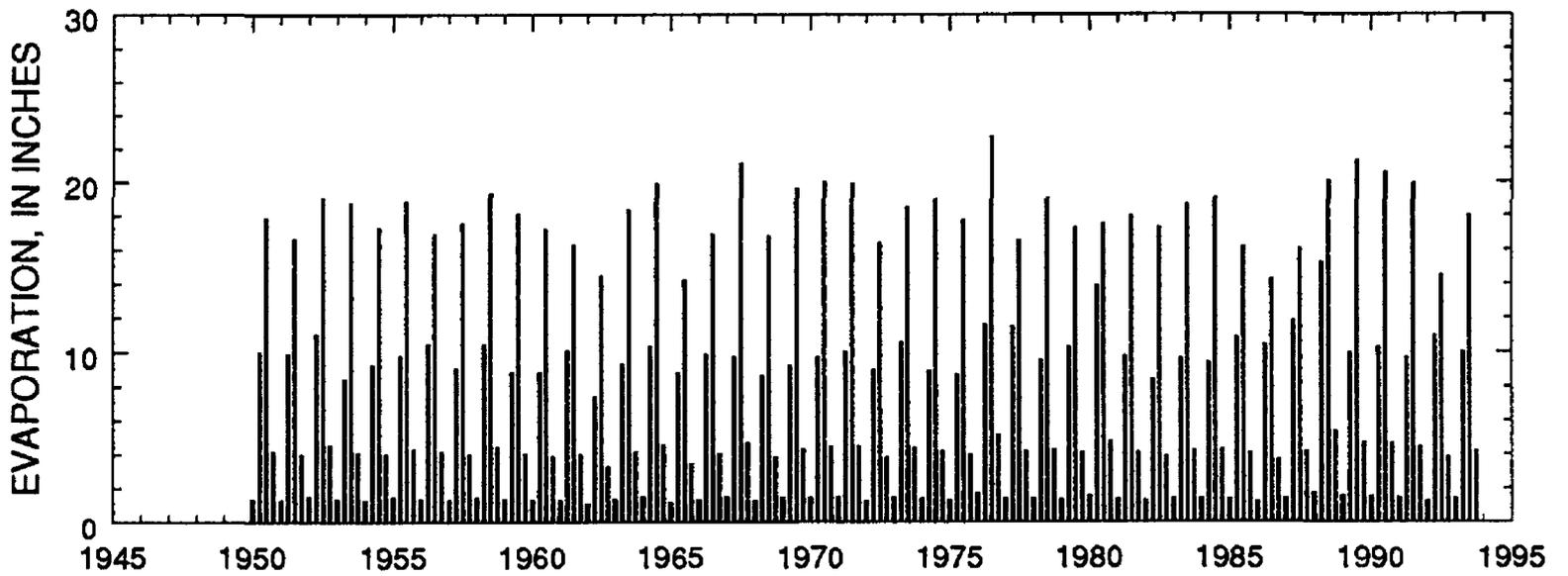
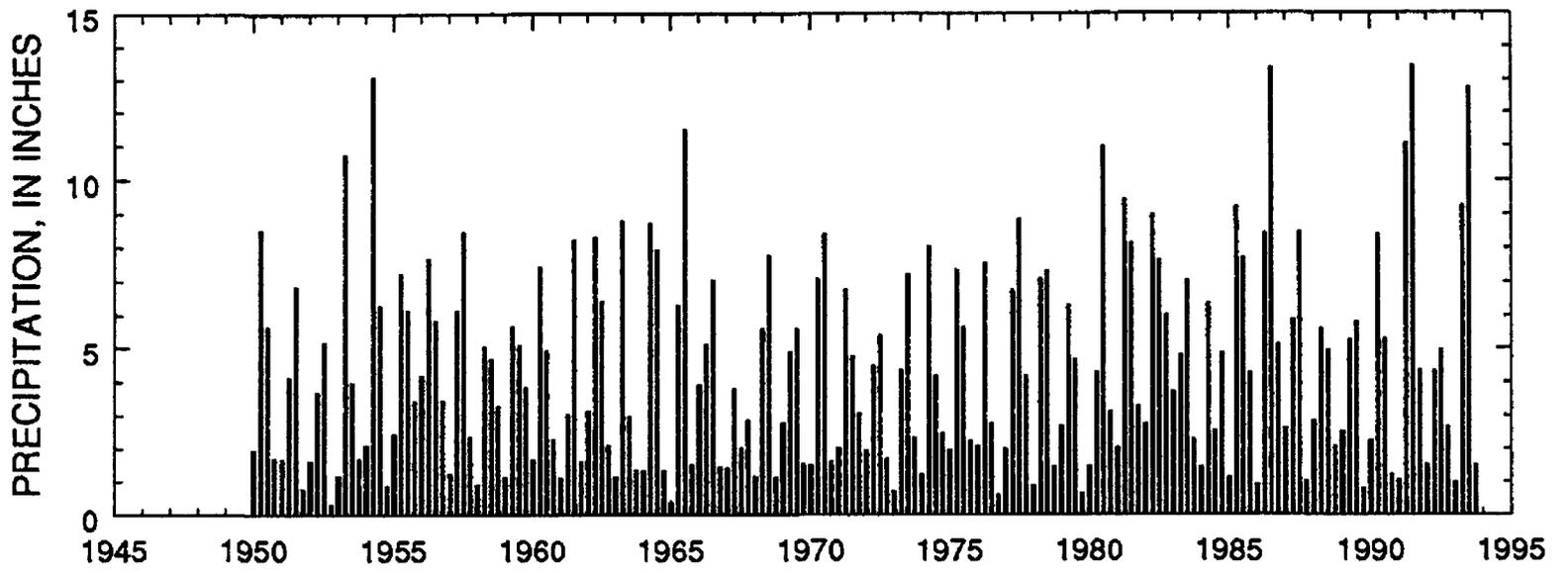


Figure 17. Quarterly precipitation, evaporation, and inflow for Devils Lake, 1950-93.

February, March), spring (April, May, June), summer (July, August, September), and fall (October, November, December). The fitted WMB model is given by

$$\Delta S_q(t) = 329.17 + 0.94 \left[\frac{1}{12} P_q(t) A_q(t) \right] - 0.97 \left[\frac{1}{12} E_q(t) A_q(t) \right] + 1.02 Q_q(t) + R(t), \quad (20)$$

where

- $\Delta S_q(t)$ is quarterly storage change, in acre-feet;
- $P_q(t)$ is quarterly precipitation falling on the lake surface, in inches;
- $A_q(t)$ is quarterly lake-surface area, in acres;
- $E_q(t)$ is quarterly evaporation from the lake surface, in inches;
- $Q_q(t)$ is quarterly total inflow, in acre-feet; and
- $R(t)$ is the quarterly model residual, in acre-feet.

Coefficients were fitted using multiple linear regression. Because monthly data were aggregated to obtain quarterly data, a nonzero intercept exists and the coefficients for precipitation, evaporation, and inflow are not exactly equal to one. Modeled and recorded storage changes [the modeled changes are obtained by setting $R(t) = 0$ in equation 20] are shown in figure 18 for 1950-93. Because of the large coefficient of determination, the residuals, $R(t)$, are set to zero. This will have a negligible smoothing effect on the lake-level traces.

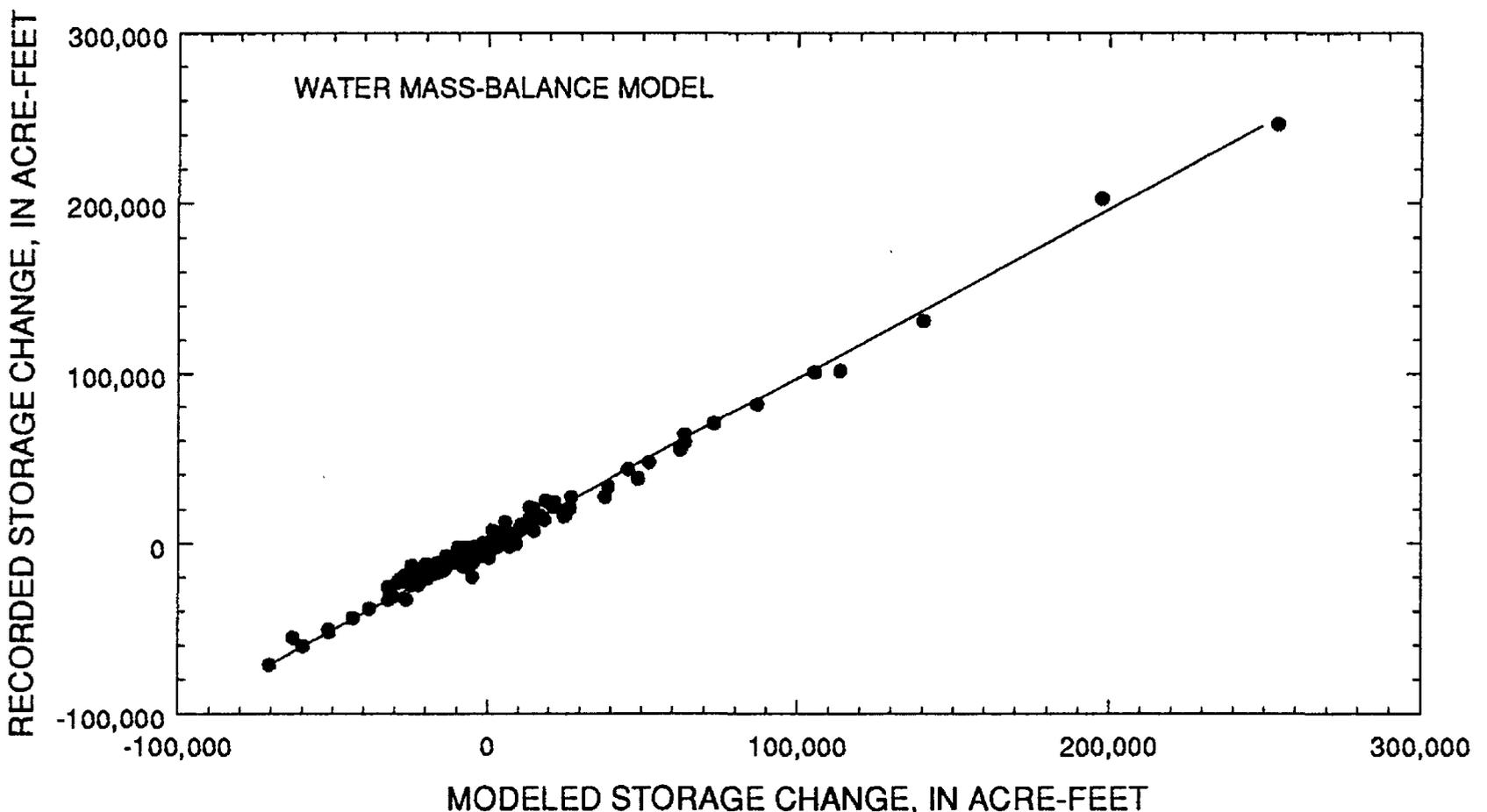


Figure 18. Modeled and recorded quarterly storage changes for Devils Lake, 1950-93.

Before the WMB model (eq. 20) can be used to generate future lake-level traces, a time-series model must be developed to generate $P_q(t)$, $E_q(t)$, and $Q_q(t)$. The subscript q will hereafter be dropped with the understanding that all inputs are quarterly time series. The model should reproduce the probability distributions of each of the inputs as well as any significant autocorrelations or crosscorrelations in the recorded data. Normality transformations were used to approximately maintain the marginal probability distributions; and a multivariate, periodic autoregressive model (Salas and others, 1985) was used to reproduce the correlation structure.

Quarterly inputs were transformed to an approximate Gaussian time series using the same power transformation (eq. 4) that was used to transform annual mean lake volumes. However, the transformation parameter was allowed to depend on the season. Because of strong temporal persistence in the annual mean lake volume time series, $X(t)$, equations 6 and 7 were used to select the transformation parameter for $X(t)$ in conjunction with the ARMA model for reproducing the persistence. However, because the quarterly inputs have no significant autocorrelation from year to year, the methods from Box and Cox (1964) were used to directly estimate the transformation parameters. For example, when considering precipitation, in all four quarters, the best transformation-parameter value from the set $\lambda \in \{1, 1/2, 1/3, 1/4, 1/5, 0\}$ was $\lambda = 1/3$. The transformed data then were standardized by subtracting the quarterly mean, $\hat{\mu}$, and dividing by the quarterly standard deviation, $\hat{\sigma}$, as follows:

$$P^*(t) = \frac{[P^{1/3}(t) - \hat{\mu}]}{\hat{\sigma}}, \quad (21)$$

where

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^n P^{1/3}(t) \text{ and } \hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n [P^{1/3}(t) - \hat{\mu}]^2. \quad (22)$$

For example, $\hat{\mu}$ is 1.88 and $\hat{\sigma}^2$ is 0.20^2 in the spring. To simplify the notation, the subtraction of 1 and the division by λ in the original Box-Cox transformation (eq. 4) was absorbed in the mean and variance for the definition of standardized precipitation, $P^*(t)$.

The power-transformation procedure was applied to precipitation data for the remaining three quarters and to evaporation and inflow data for all four quarters. Results of the power transformations are

$$\text{Winter} \quad P^*(t) = \frac{P(t)^{1/3} - 1.19}{0.19} \quad (23)$$

$$E^*(t) = \frac{E(t) - 1.43}{0.24} \quad (24)$$

$$Q^*(t) = \frac{Q(t) - 6,051.5}{5,102} \quad (25)$$

$$\text{Spring} \quad P^*(t) = \frac{P(t)^{1/3} - 1.88}{0.20} \quad (26)$$

$$E^*(t) = \frac{\ln \{E(t) - 5.1\} - 1.57}{0.25} \quad (27)$$

$$Q^*(t) = \frac{[Q(t) - 4,200]^{1/5} - 7.03}{2.13} \quad (28)$$

$$\text{Summer} \quad P^*(t) = \frac{P(t)^{1/3} - 1.85}{0.25} \quad (29)$$

$$E^*(t) = \frac{E(t) - 18.02}{1.87} \quad (30)$$

$$Q^*(t) = \frac{\ln \{Q(t) + 3,000\} - 9.81}{0.82} \quad (31)$$

$$\text{Fall} \quad P^*(t) = \frac{P(t)^{1/3} - 1.27}{0.25} \quad (32)$$

$$E^*(t) = \frac{E(t) - 4.22}{0.37} \quad (33)$$

$$Q^*(t) = \frac{\ln \{Q(t) + 13,000\} - 9.57}{0.47} \quad (34)$$

The spring evaporation value and the spring, summer, and fall inflow values were modified by adding a constant before applying the power-transformation procedure. Because the standard Box-Cox transformed data in these cases were not close enough to a normal distribution, a location parameter was included in addition to the power-transformation parameter. The location parameter was estimated as outlined in Box and Cox (1964).

A time-series model then was fitted to the transformed, standardized values for precipitation, evaporation, and inflow in order to maintain the autocorrelation and crosscorrelation structure. This was complicated by the fact that the correlations depend on season as well as lag. For example, the correlation between spring inflow and the previous winter's precipitation is 0.47 and the correlation between fall inflow and the previous summer's precipitation is 0.03 even though the lag in both cases is one season.

In order to maintain the complicated seasonal crosscorrelation structure in the inputs, a multivariate, periodic autoregressive model was fitted to the standardized inputs using model fitting procedures described in supplement 1. The fitted model is

$$\text{Winter} \quad P^*(t) = X(t) \quad (35)$$

$$E^*(t) = Y(t) \quad (36)$$

$$Q^*(t) = 0.47P^*(t-1) + 0.89Z(t) \quad (37)$$

$$\text{Spring} \quad P^*(t) = X(t) \quad (38)$$

$$E^*(t) = 0.48E^*(t-1) + 0.31P^*(t-4) - 0.30P^*(t) + 0.83Y(t) \quad (39)$$

$$Q^*(t) = 0.46P^*(t-1) + 0.32Q^*(t-2) + 0.21Q^*(t-4) + 0.78Z(t) \quad (40)$$

$$\text{Summer} \quad P^* (t) = X (t) \quad (41)$$

$$E^* (t) = 0.49E^* (t-2) - 0.46P^* (t) + 0.74Y (t) \quad (42)$$

$$Q^* (t) = 0.54Q^* (t-1) + 0.46E^* (t) + 0.34P^* (t) + 0.73Z (t) \quad (43)$$

$$\text{Fall} \quad P^* (t) = X (t) \quad (44)$$

$$E^* (t) = 0.73E^* (t-1) + 0.52E^* (t-2) + 0.13Y (t) \quad (45)$$

$$Q^* (t) = 0.50Q^* (t-1) + 0.32P^* (t) + 0.80Z (t) , \quad (46)$$

where

$X (t)$, $Y (t)$, and $Z (t)$ are mutually independent, standard normal, white noise time series.

Each of the coefficients in the model is significantly different from zero at the 5-percent significance level. Many of the coefficients make sense from a hydrologic standpoint, but other coefficients are not as easily interpreted. For example, the negative coefficients relating evaporation to concurrent precipitation (eqs. 39 and 42) and the positive coefficients relating inflow to concurrent or lagged precipitation (eqs. 37, 40, 43, and 46) make sense. The coefficients relating summer and fall inflows to the previous seasons' inflows (eqs. 43 and 46) also make sense. The lagged autoregressive coefficient relating spring evaporation to the previous spring's precipitation (eq. 39) probably is spurious but should not have any significant bearing on the model output. The coefficients relating evaporation to lagged values of itself (eqs. 39, 42, and 45) probably are attributable, in part, to true persistence in the evaporation time series and, in part, to the way in which the lake-evaporation data were derived.

Spring inflow (eq. 40) had the largest effect on the lake-level frequency analysis. The spring inflow from one year is related to the spring and fall inflows from the previous year. This persistence in inflow on an interannual scale and the fact that extremely large lake-level rises can occur in conjunction with large spring inflow has a significant effect on the WMB model output.

Analysis of Model Output

The first step in generating the lake-level time series using the WMB model is to generate the seasonal precipitation, evaporation, and inflow time series. As with the ALV model, either conditional or unconditional simulations can be used. For conditional simulations, the required starting values of the inputs are replaced by recorded values. For example, if conditional simulations are used beginning with winter 1994, the required values before 1994 are replaced by recorded values. The only starting values required by the model (eqs. 35 to 46) would be spring and fall 1993 precipitation and spring and fall 1993 inflow. If, however, the conditional simulations are used beginning with summer 1994, the required starting values would be winter and spring 1994 evaporation and spring 1994 inflow.

For unconditional simulations, the required starting values of the standardized inputs are selected randomly from their stationary, joint probability distribution. The model can be used to derive an exact multivariate, normal distribution for the required starting values. Alternatively, approximate unconditional simulations can be obtained by setting all initial values of the standardized inputs to zero, generating a long

conditional simulation, and then dropping the first part of the conditional simulation. The initial conditions are assumed to have a negligible effect after 50 years because of the relatively weak dependence in the inputs.

The algorithm for generating conditional simulations of seasonal precipitation, evaporation, and inflow is as follows:

1. Generate independent, standard normal, white noise time-series variables, $X(t)$, $Y(t)$, and $Z(t)$, using a normal, random number generator.
2. Compute $P^*(t)$, $E^*(t)$, and $Q^*(t)$ from the multivariate, periodic autoregressive model (eqs. 35 to 46). Initial conditions are obtained as described earlier.
3. Reverse the power transformations (eqs. 23 to 34) to obtain $P(t)$, $E(t)$, and $Q(t)$ in original units.

To compare the model output to the recorded seasonal inputs for 1950-93, 100 44-year unconditional simulations were obtained from the model. Statistics of the recorded data and corresponding model values are given in table 7. The model closely reproduces the statistics of the recorded seasonal precipitation. It also closely reproduces the statistics of the recorded seasonal evaporation except for the spring skewness coefficient and the lag-1 autocorrelation coefficient. This discrepancy has a negligible effect on the model results because of the small variability of spring evaporation.

Table 7. Selected statistics of precipitation, evaporation, and inflow data for 1950-93 and corresponding values generated from the water mass-balance model

{Number in parentheses is average of 100 values generated from the model; numbers in brackets represent range that included 90 percent of the values produced from the model}

Season	Mean	Standard devistion	Skewness coefficient	Lag-1 sutocorrelation coefficient
	Precipitation (inches)			
Winter	1.83 (1.81) [1.61 to 2.04]	0.86 (0.84) [0.68 to 1.05]	0.87 (0.84) [0.29 to 1.51]	0.07 (0.02) [-0.25 to 0.34]
Spring	6.84 (6.88) [6.38 to 7.37]	2.16 (2.11) [1.73 to 2.50]	0.53 (0.55) [0.05 to 1.14]	-0.05 (-0.01) [-0.24 to 0.25]
Summer	6.70 (6.74) [6.13 to 7.38]	2.65 (2.60) [2.07 to 3.17]	0.81 (0.67) [0.14 to 1.45]	0.18 (-0.01) [-0.30 to 0.27]
Fall	2.31 (2.29) [1.93 to 2.65]	1.31 (1.26) [0.93 to 1.65]	0.81 (0.92) [0.37 to 1.84]	0.21 (-0.01) [-0.30 to 0.27]
Annual	17.69 (17.72) [16.95 to 18.59]	4.16 (3.67) [3.13 to 4.45]	0.84 (0.36) [-0.12 to 0.92]	0.09 (-0.04) [-0.28 to 0.19]

Table 7. Selected statistics of precipitation, evaporation, and inflow data for 1950-93 and corresponding values generated from the water mass-balance model—Continued

[Number in parentheses is average of 100 values generated from the model; numbers in brackets represent range that included 90 percent of the values produced from the model]

Season	Mean	Standard deviation	Skewness coefficient	Lag-1 autocorrelation coefficient
Evaporation (inches)				
Winter	1.38 (1.43) [1.37 to 1.49]	0.12 (0.23) [0.19 to 0.28]	0.45 (-0.04) [-0.53 to 0.40]	0.05 (-0.01) [-0.25 to 0.24]
Spring	9.98 (9.98) [9.71 to 10.30]	1.35 (1.29) [0.95 to 1.62]	1.71 (0.64) [0.09 to 1.41]	0.68 (0.45) [0.27 to 0.63]
Summer	18.02 (18.03) [17.67 to 18.43]	1.87 (1.82) [1.48 to 2.17]	-0.03 (-0.02) [-0.56 to 0.53]	0.18 (0.22) [-0.05 to 0.44]
Fall	4.22 (4.22) [4.14 to 4.32]	0.37 (0.36) [0.30 to 0.43]	0.45 (0.02) [-0.43 to 0.54]	0.84 (0.84) [0.73 to 0.90]
Annual	33.61 (33.66) [33.02 to 34.45]	3.00 (2.96) [2.42 to 3.55]	0.44 (0.15) [-0.26 to 0.72]	0.04 (-0.03) [0.28 to 0.18]
Inflow (acre-feet)				
Winter	6,050 (6,059) [4,994 to 7,060]	5,100 (4,962) [4,110 to 5,704]	0.57 (0.07) [-0.46 to 0.61]	0.18 (0.14) [-0.09 to 0.37]
Spring	39,100 (39,911) [25,760 to 57,407]	48,400 (48,193) [24,327 to 83,528]	2.64 (2.40) [1.28 to 4.06]	0.10 (0.06) [-0.17 to 0.29]
Summer	17,700 (17,333) [11,451 to 24,731]	35,600 (23,755) [15,274 to 35,577]	4.69 (2.06) [0.94 to 3.72]	0.32 (0.45) [0.15 to 0.72]
Fall	3,100 (3,184) [1,058 to 5,421]	9,200 (7,641) [5,663 to 10,896]	3.16 (1.13) [0.45 to 2.09]	0.82 (0.43) [0.21 to 0.66]
Annual	65,500 (66,487) [47,010 to 87,261]	73,300 (66,503) [38,990 to 103,456]	2.13 (1.97) [0.96 to 3.45]	0.06 (0.22) [0.02 to 0.52]

Winter and spring inflows show no significant deviations between the recorded data and the model values. However, the recorded summer standard deviation and skewness coefficient and the recorded fall skewness coefficient and lag-1 autocorrelation coefficient all are outside the 90-percent range of values produced by the model. The recorded summer and fall inflows for 1993 were the largest on record (fig. 19). The large inflow greatly affected the statistics of the recorded data, especially those for the standard deviation, skewness coefficient, and lag-1 autocorrelation coefficient. However, the Box-Cox transformed inflows indicate the large summer and fall inflows were not as extreme as indicated by the

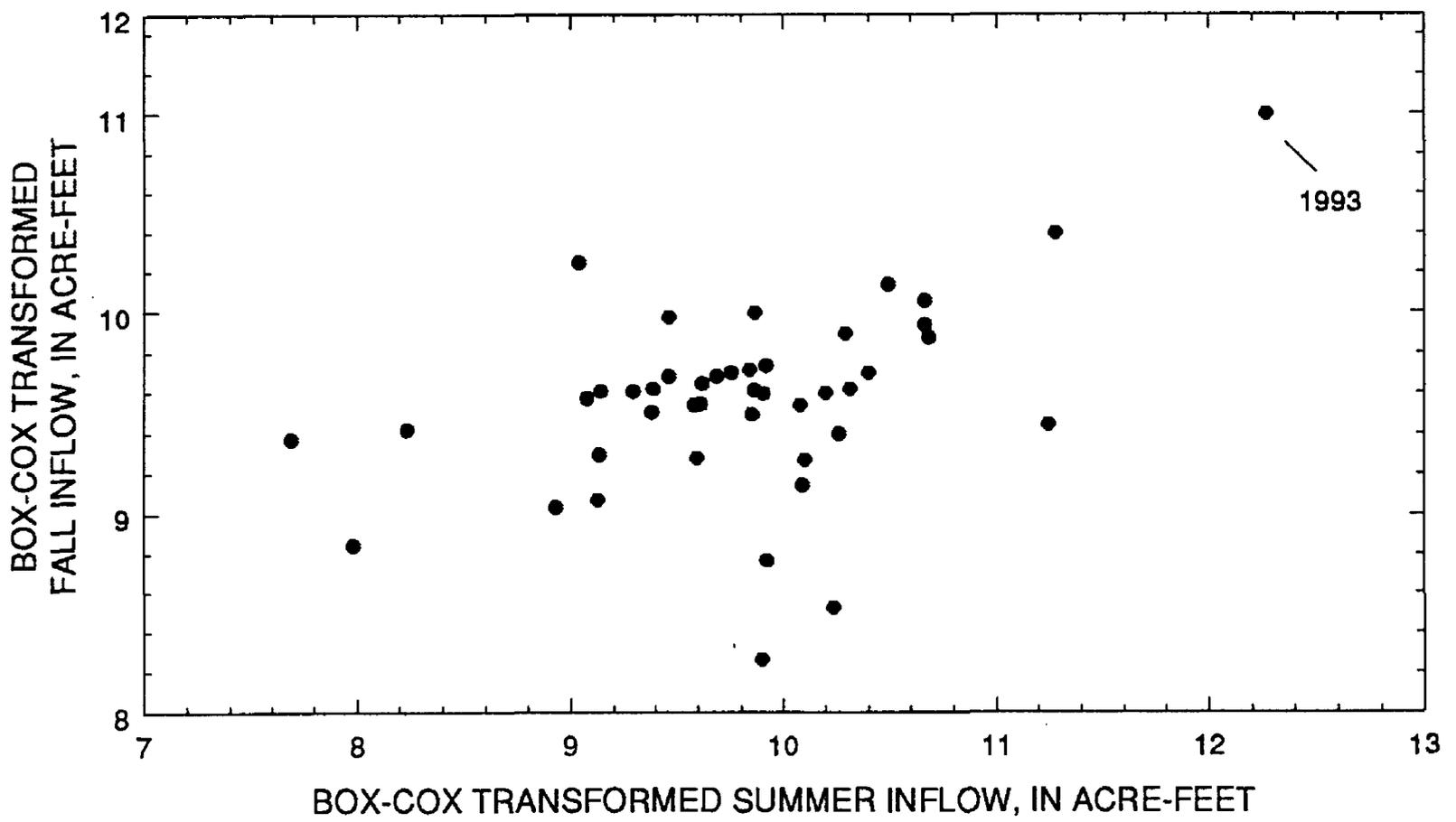
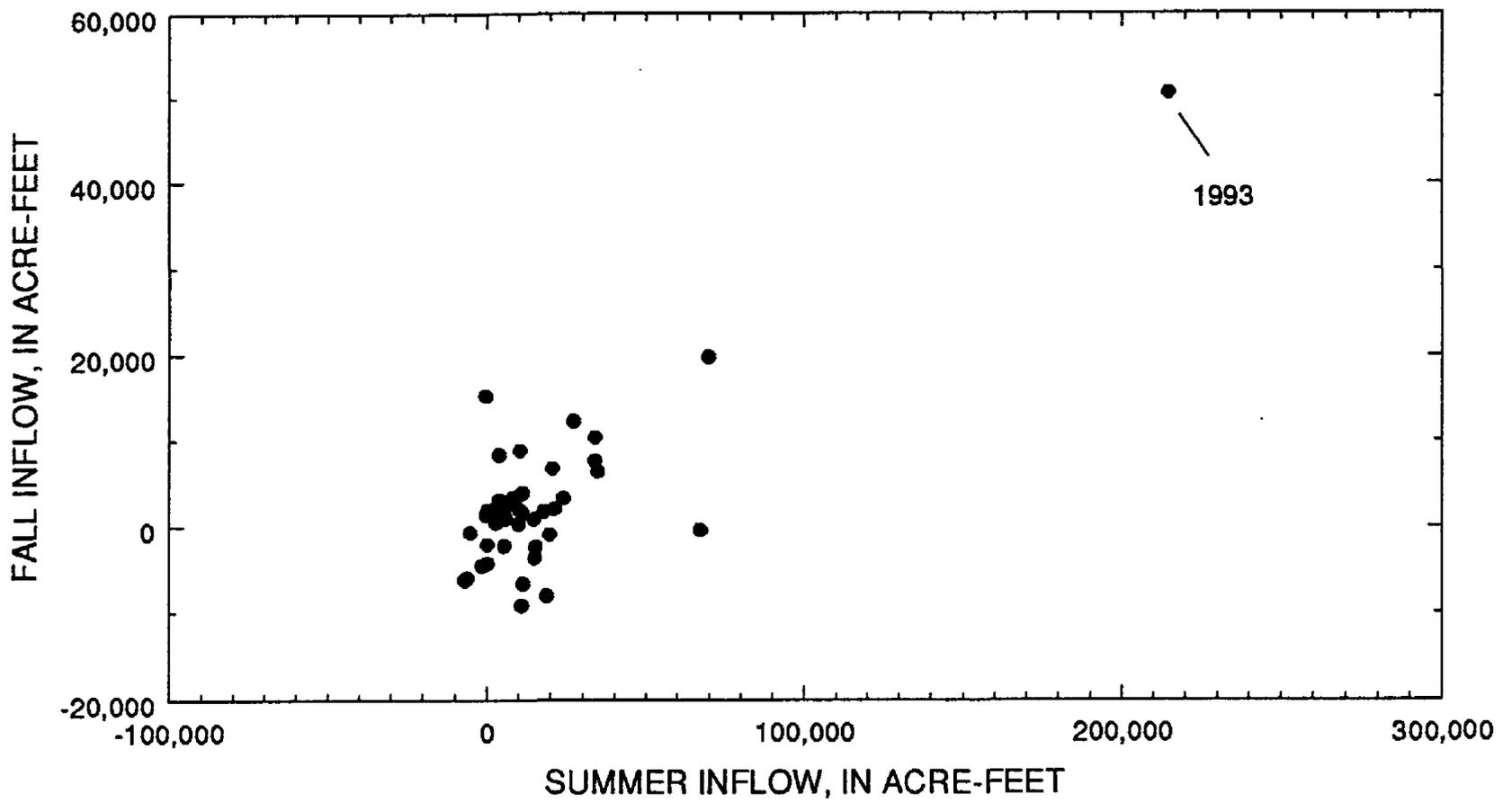


Figure 19. Relation of summer inflow to fall inflow for Devils Lake, 1950-93.

untransformed recorded data. The variable power transformations are designed to alleviate this type of situation. Because the effect of the extremes is reduced, the model values probably are a good representation of the true parameters.

Conditional simulations of annual maximum lake levels can be obtained as follows:

1. Generate time series of $P(t)$, $E(t)$, and $Q(t)$ as described in steps 1 through 3 for generating conditional simulations of seasonal precipitation, evaporation, and inflow.
2. Given the current lake volume and lake-surface area, use the fitted WMB model (eq. 20) to generate a time series of seasonal lake volumes [$R(t)$ was set to zero for the generated data].
3. Convert lake volumes to lake levels using the Devils Lake elevation-capacity table.
4. Obtain a time series of annual maximum lake levels by using the largest of the four seasonal lake levels in each year.

This procedure was used to generate a single, 400-year, unconditional simulation lake-level trace from the fitted WMB model (eq. 20). To approximate an unconditional simulation trace, the starting lake volume and lake-surface area were set to the fall 1993 recorded values, and the first 200 years of the conditional simulation trace were dropped. The generated lake-level trace is shown in figure 20. The corresponding time series of lake-level changes also is shown in figure 20.

To determine statistics of the simulated lake-level changes and to compare output from the WMB model and the ALV model (table 2), 100 93-year unconditional simulation lake-level traces were generated from the WMB model. The statistics of the simulated lake-level changes are given in table 8. The WMB model produces fewer large lake-level declines than the ALV model, as indicated in figures 9 and 20. Therefore, the skewness coefficient is more closely reproduced by the WMB model (table 8) than by the ALV model (table 2). However, the WMB model tends to produce a smaller standard deviation than the recorded lake-level changes.

Table 8. Selected statistics of recorded lake-level changes for 1901-93 and corresponding values generated from the water mass-balance model

	Recorded lake-level change ¹	Water mass-balance model	
		Average simulated lake-level change ²	90-percent range ³
Mean (feet)	-0.041	-0.031	-0.129 to 0.080
Standard deviation (feet)	1.61	1.26	0.99 to 1.69
Skewness coefficient	1.91	1.87	0.81 to 3.74
Lag-1 autocorrelation coefficient	.27	.27	0.10 to 0.45
Lag-2 autocorrelation coefficient	.01	.03	-0.09 to 0.17

¹Difference between maximum lake level in year t and maximum lake level in year $t-1$.

²Based on 100 93-year unconditional simulation lake-level traces generated from the water mass-balance model.

³Numbers represent range that included 90 percent of the values produced from the model.

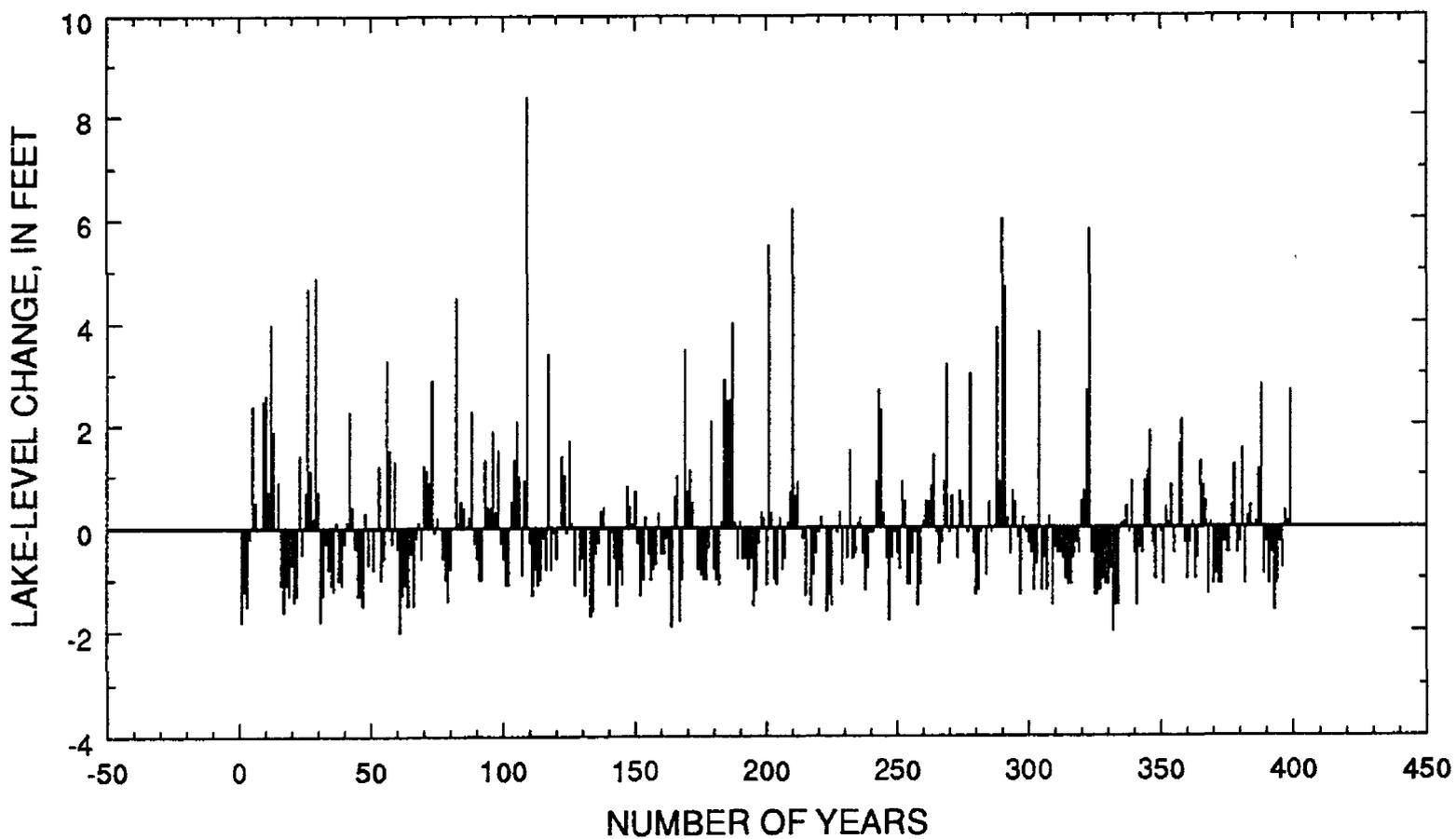
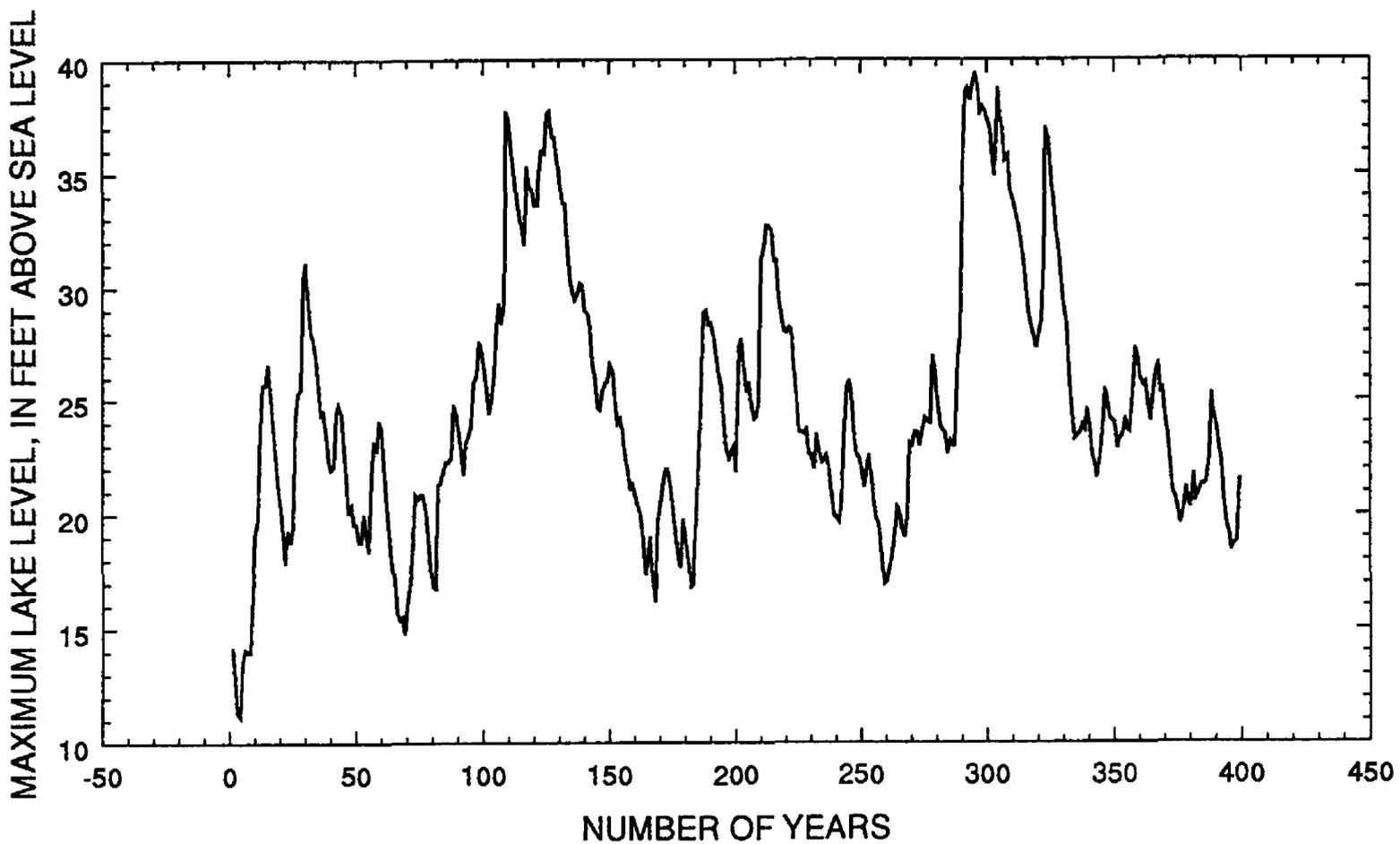


Figure 20. Unconditional simulation lake-level trace from the water mass-balance model and corresponding lake-level changes for Devils Lake.

To demonstrate the effects of conditioning on lake-level frequency analysis, 2,000 conditional simulation lake-level traces, each 50 years in length, were generated from the statistical WMB model. Conditional simulation exceedance levels for exceedance probabilities of 0.01, 0.25, 0.50, 0.75, and 0.99 are shown in figure 21. The effect of the conditioning is shown at the beginning of the simulation period and persists for several decades. The exceedance levels at the end of the simulation period closely approximate the stationary exceedance levels of the model. A comparison of these stationary values with the stationary values from the ALV model (fig. 10; table 3) indicates that the upper exceedance levels (25 percent and above) agree fairly well but the lower levels do not.

Parameter-Uncertainty Analysis

The effect of parameter uncertainty in the fitted WMB model must be included to determine the reliability of the exceedance levels from the model. This is especially important considering the relatively short calibration period (1950-93).

A parameter-uncertainty distribution that takes into account sampling error in parameter estimates is discussed in supplement 1. The parameter estimates in the uncertainty analysis include the sample means and standard deviations of the Box-Cox transformed quarterly inflows (eqs. 23 to 34; 12 means and 12 standard deviations) and the autoregressive parameters in the multivariate, periodic autoregressive model (eqs. 35 to 46; 16 autoregressive parameters). The fact that the 40 parameter estimates are not statistically independent of one another complicates the uncertainty analysis.

Conditional simulation exceedance levels for exceedance probabilities of 0.01, 0.25, 0.50, 0.75, and 0.99 are shown in figure 21. The exceedance levels were computed from 2,000 50-year traces from the WMB model using parameters adjusted for uncertainty. The parameters for each trace were sampled randomly from the parameter-uncertainty distribution rather than held fixed at their fitted values. The upper exceedance levels are particularly sensitive as indicated by a comparison of the 1-percent exceedance levels (0.01 exceedance probabilities) shown in figure 21. For example, the 1-percent exceedance level from the WMB model that includes parameter uncertainty is 4.9 feet higher in year 50 than the 1-percent exceedance level from the WMB model that has fixed parameters (table 3). Also, large discrepancies exist between the WMB exceedance levels shown in figure 21 and the ALV exceedance levels shown in figure 10. If both models are reasonable representations of the true long-term lake levels, the stationary distributions should be similar. Even though the ALV model produces traces that are not as realistic in terms of the lake hydrology as the WMB model traces, the stationary distribution of the ALV traces should be fairly accurate. The long period of record used for the ALV model results in relatively small uncertainty in the model output as shown by a comparison of the ALV output using fixed parameters and the ALV output using parameters adjusted for uncertainty (fig. 10).

The sensitivity in the upper exceedance levels from the WMB model is caused almost entirely by uncertainty in the fitted probability distributions of the quarterly inflows. The uncertainty in the precipitation and evaporation distributions was only a minor contributor. Therefore, a technique for reducing parameter uncertainty in the seasonal inflow distributions will be presented.

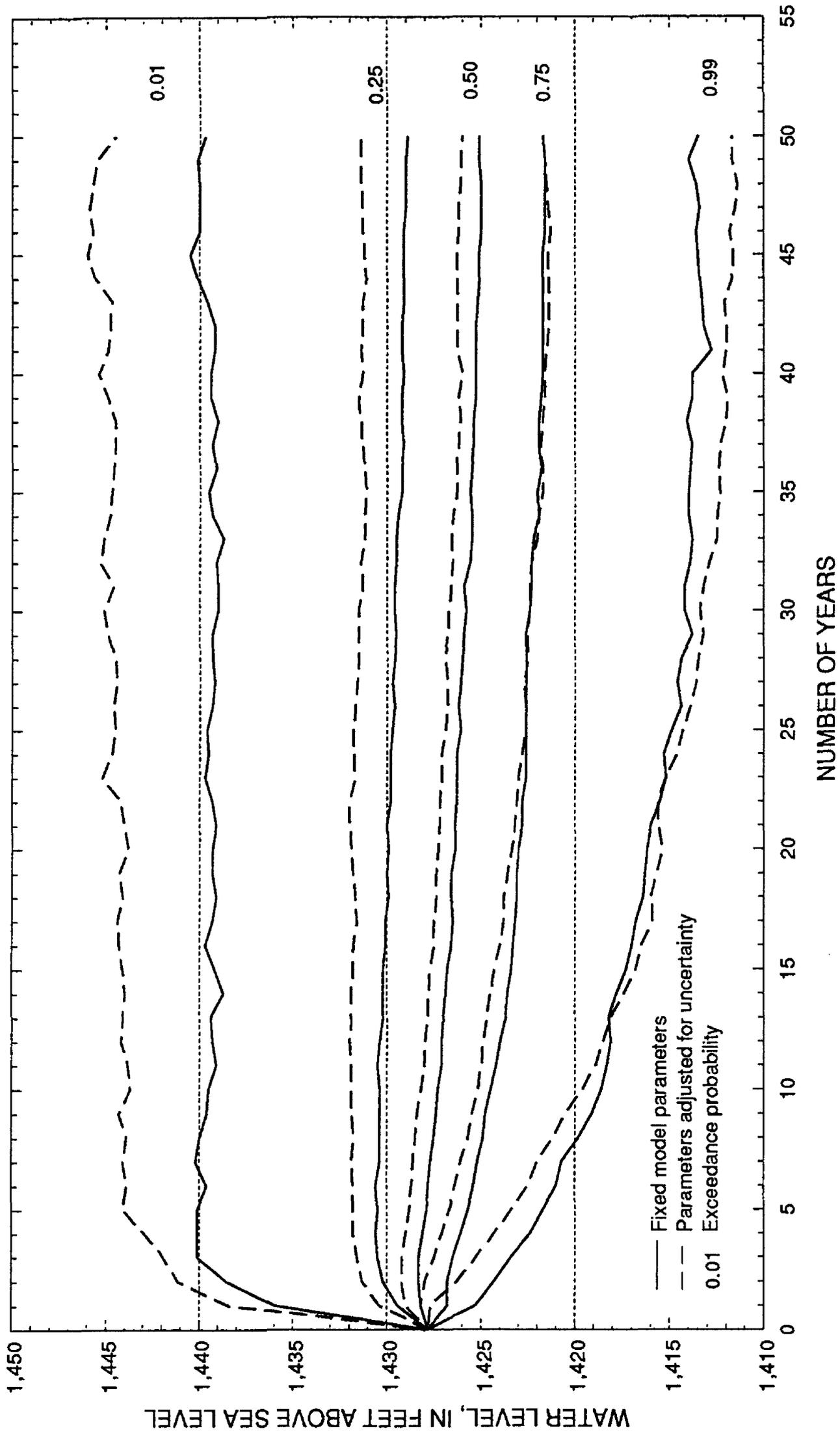


Figure 21. Exceedance levels for Devils Lake computed from 2,000 50-year traces from the water mass-balance model using fixed parameters and parameters adjusted for uncertainty. (Year 0 corresponds to 1993.)

Reduction of Parameter Uncertainty in Seasonal Inflow Distributions

The marginal probability distributions of the seasonal inflows are governed entirely by the variable power transformations (eqs. 25, 28, 31, and 34). As an example, consider spring inflow. Let

$$Y(t) = [Q(t) - 4,200]^{1/5}, \quad (47)$$

where

$Y(t)$ is the Box-Cox transformed spring inflow for year t (t is used in this section to denote years rather than quarters).

The fitted probability distribution for $Y(t)$ is a normal distribution with fitted mean,

$$\hat{\mu} = \frac{1}{44} \sum_{t=1950}^{1993} Y(t) = 7.03, \quad (48)$$

and fitted variance,

$$\hat{\sigma}^2 = \frac{1}{44} \sum_{t=1950}^{1993} [Y(t) - \hat{\mu}]^2 = (2.13)^2. \quad (49)$$

Considering sampling uncertainty in the fitted mean and assuming the multivariate, periodic autoregressive model (eqs. 35 to 46) is correct,

$$\text{Corr}[Y(t), Y(t-k)] = \phi_Y^k, \quad (50)$$

where

k is the lag, in years; and
 ϕ_Y is the lag-1 autocorrelation coefficient.

Therefore, according to Brockwell and Davis (1987, ch. 7), $\hat{\mu}$ is approximately normally distributed with an expected value and variance as follows:

$$E(\hat{\mu}) = \mu \text{ and } \text{Var}(\hat{\mu}) = \frac{1 + \phi_Y \sigma_Y^2}{1 - \phi_Y} \frac{1}{n} \quad (51)$$

where

σ_Y^2 is the variance of $Y(t)$, and
 n is the sample size for 1950-93 (44 years).

An alternative method for determining the mean of $Y(t)$ was designed to reduce the variance in the estimated mean. The new method uses long-term streamflow data for the Red River of the North at Grand Forks to augment the inflow data for Devils Lake for 1950-93. The Red River streamflow data, which were selected because of the long period of record, are correlated with the Devils Lake inflow data. The natural log-transformed spring streamflow data (average daily streamflow for April, May, and June, in cubic feet per second) for the Red River of the North at Grand Forks are plotted in figure 22. The mean of the recorded data for 1883-1949 (when Devils Lake inflow data are not available) and the mean of the

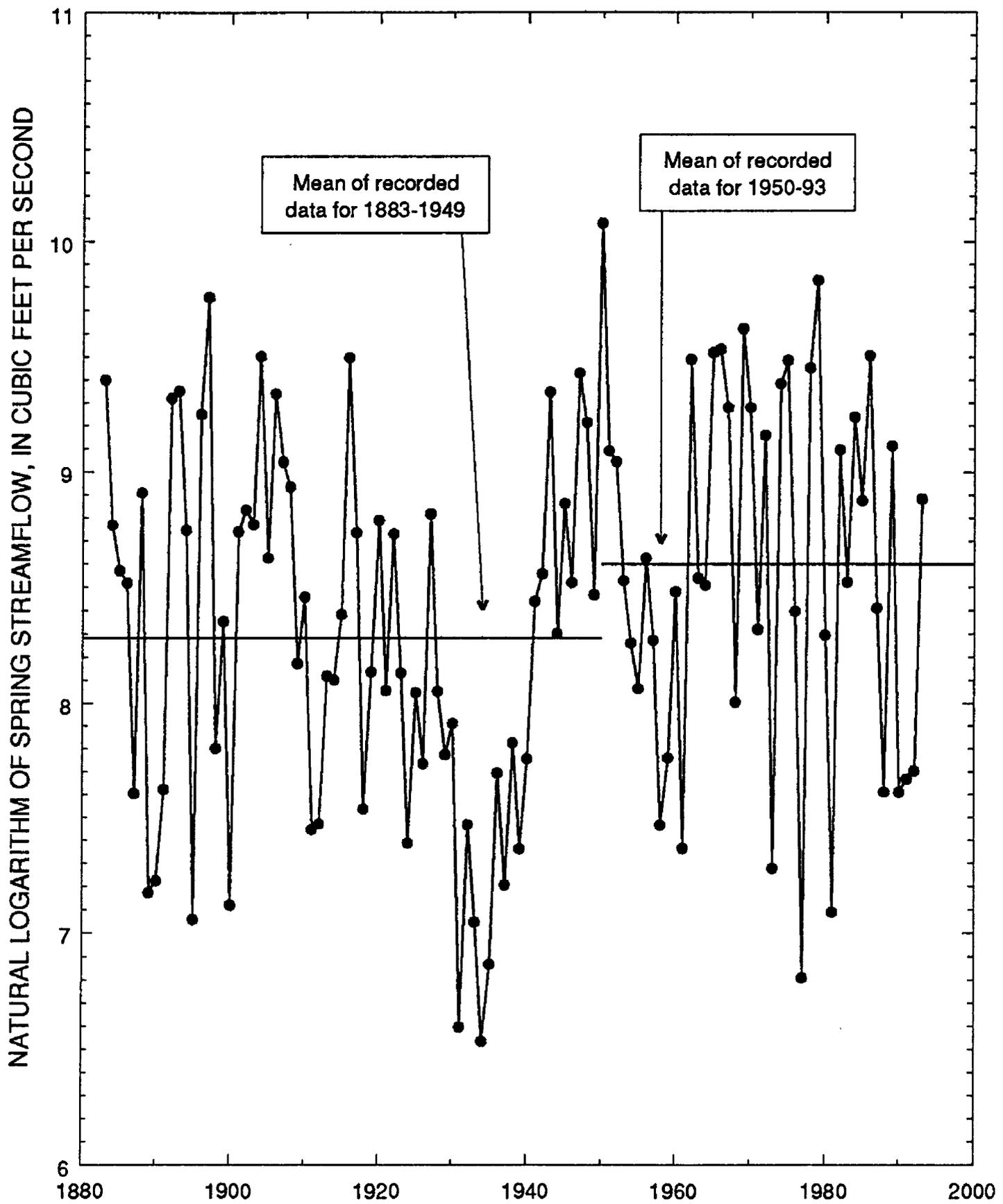


Figure 22. Spring streamflow for the Red River of the North at Grand Forks, 1883-1993.

recorded data for 1950-93 (when concurrent Devils Lake inflow data are available) also are shown. The relation between the log-transformed spring streamflow data for the Red River of the North at Grand Forks and the Box-Cox transformed spring inflow data for Devils Lake for 1950-93 is shown in figure 23. The correlation coefficient is 0.49.

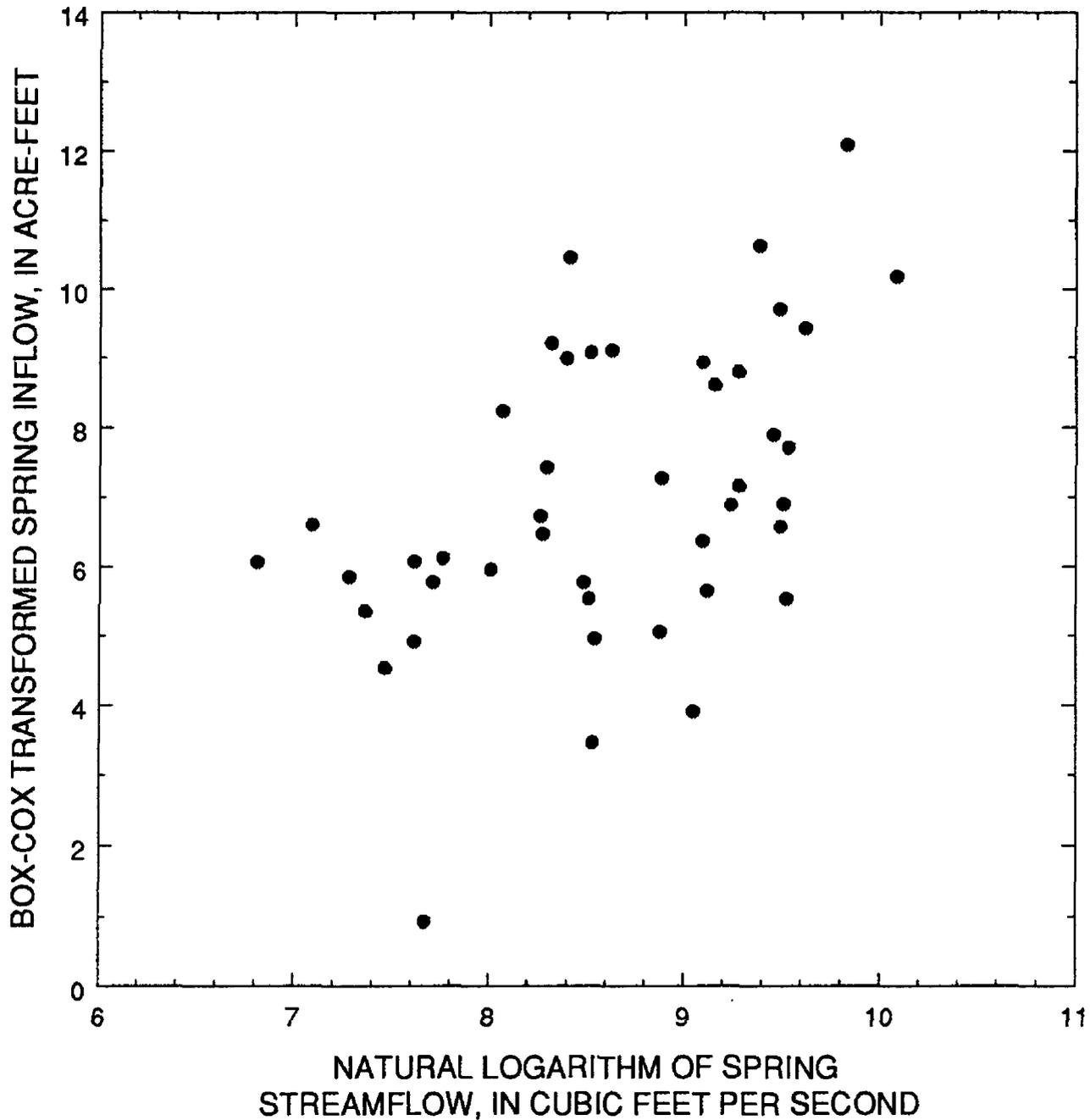


Figure 23. Relation of spring streamflow for the Red River of the North at Grand Forks to spring inflow for Devils Lake, 1950-93.

Let $X(t)$ denote the natural log-transformed spring streamflow data for the Red River of the North at Grand Forks for year t . To use the extra information available through $X(t)$, the following method for determining the mean of $Y(t)$ was considered:

$$\hat{\mu}' = \hat{\mu} + a(\bar{X}_1 - \bar{X}_2), \quad (52)$$

where

$\hat{\mu}$ is given in equation 48,

a is a constant, and

$$\bar{X}_1 = \frac{1}{67} \sum_{t=1883}^{1949} X(t) \text{ and } \bar{X}_2 = \frac{1}{44} \sum_{t=1950}^{1993} X(t). \quad (53)$$

Assuming that $Y(t)$ and $X(t)$ are both stationary time series, equation 52 is clearly an unbiased estimator of the true mean of $Y(t)$ for any value of the constant a . Assuming the autocorrelation structure (eq. 50) for $Y(t)$ and a similar assumption for $X(t)$, the value of a that minimizes the variance of equation 52 is shown in supplement 1 to be approximated by

$$a_m = \frac{n_1}{n_1 + n_2} \frac{\sigma_Y}{\sigma_X} \left\{ \frac{(1 + \phi_Y)(1 - \phi_X)}{(1 - \phi_Y)(1 + \phi_X)} \right\}^{1/2} p, \quad (54)$$

where

ϕ_X is the lag-1 autocorrelation coefficient for $X(t)$;

n_1 is the sample size for 1883-1949 (67 years);

n_2 is the sample size for 1950-93 (44 years); and

$$p = \text{Corr}[Y(t) - \phi_Y Y(t-1), X(t) - \phi_X X(t-1)]. \quad (55)$$

The approximate variance of the resulting adjusted sample mean (see supplement 1) is

$$\text{Var}[\hat{\mu} + a_m(\bar{X}_1 - \bar{X}_2)] = \frac{\sigma_Y^2}{n_2} \frac{1 + \phi_Y}{1 - \phi_Y} \left\{ 1 - p^2 \frac{n_1}{n_1 + n_2} \right\}. \quad (56)$$

The same procedure was used to obtain adjusted estimates of the mean of transformed inflows for winter, summer, and fall (table 9). Estimates of the correlation parameters obtained from the recorded data also are given in table 9. The largest relative adjustment occurs in the spring when the crosscorrelation parameter was the largest ($p = 0.50$). The adjustments for winter, summer, and fall are very small and result in only slight decreases in the standard deviations of the fitted means.

In order to illustrate the effect of the adjustment, 2,000 conditional simulation lake-level traces were generated using the adjusted mean, $\hat{\mu}'$, rather than the original mean, $\hat{\mu}$. The effect of the adjustment is shown in figure 24. The adjustment essentially lowered the stationary distribution of the lake levels by about 2 feet (table 3). This downward adjustment results in much closer agreement between the stationary exceedance levels of the ALV model (fig. 10) and those of the WMB model (fig. 24), especially in the upper levels. Some disparity exists in the lower levels, but, when considered in terms of lake volumes rather than lake levels, the disparity is not large. The lower exceedance levels of the ALV model probably are caused by the model's propensity to give unrealistically large lake-level declines when the lake is already low, as discussed previously. For this reason, the lower exceedance levels of the WMB model probably are more realistic than those of the ALV model.

Table 9. Estimated means of transformed quarterly inflows before and after adjustment to reduce parameter uncertainty

	Before adjustment		After adjustment		Correlation parameters		
	Mean	Standard deviation of the mean	Mean	Standard deviation of the mean	Lag-1 correlation coefficient for X	Lag-1 correlation coefficient for Y	Cross-correlation coefficient, ρ
Winter	¹ 6,051.50	¹ 1,143.40	¹ 5,710.30	¹ 1,117.60	0.54	0.14	0.26
Spring	7.03	.47	6.73	.43	.30	.37	.50
Summer	9.81	.18	9.78	.18	.53	.34	.17
Fall	9.57	.01	9.55	.01	.70	.22	.22

¹Value was not transformed.

Further comparison indicates the upper exceedance levels of the WMB model (fig. 24) increase much more rapidly than those of the ALV model (fig. 10). For example, the 1-percent exceedance level for the WMB model that has parameters adjusted for uncertainty and reduced on the basis of streamflow for the Red River of the North at Grand Forks increases 15.6 feet from year zero to year 5; however, the 1-percent exceedance level for the ALV model increases 8.0 feet (table 3). Although this does not signify a problem with either of the models, it does indicate that care must be taken in the interpretation of lake-level frequency distributions. The WMB model, which is based on quarterly precipitation, evaporation, and inflow, uses more data than the ALV model, which is based only on annual maximum and minimum lake volumes. Because the WMB model uses more data, the conditional simulation exceedance levels are more sensitive to antecedent conditions. Record precipitation and inflow in the summer and fall of 1993 are largely responsible for the rapid increase in the upper exceedance levels of the WMB model (fig. 24).

To better illustrate the effect of the conditioning on the WMB model output, an additional 2,000 traces were generated from the WMB model. The traces were conditioned only on the starting lake level. Starting values of precipitation, evaporation, and inflow for 1993 were generated from stationary, joint probability distributions rather than held fixed at recorded values. The early exceedance levels are extremely sensitive to antecedent conditions as indicated by comparing the 1-percent exceedance levels in figures 24 and 25. Exceedance levels shown in figure 25 for years 1 to 20 are more comparable to exceedance levels of the ALV model than those shown in figure 24 for years 1 to 20.

SUMMARY AND CONCLUSIONS

Two approaches were used to estimate future lake-level probabilities of Devils Lake. The first approach is based on an annual lake-volume (ALV) model, and the second approach is based on a statistical water mass-balance (WMB) model that generates seasonal lake volumes on the basis of seasonal precipitation, evaporation, and inflow. A multivariate time-series model was developed to generate future realizations of seasonal precipitation, evaporation, and inflow, which are used in conjunction with the WMB model to generate future lake-level traces.

The ALV model is based on a decomposition of the annual maximum lake volume into the sum of two parts. The first part, $X(t)$, is the annual mean lake volume; and the second part, $D(t)$, is the difference between the annual maximum lake volume and the annual mean lake volume. Autoregressive moving

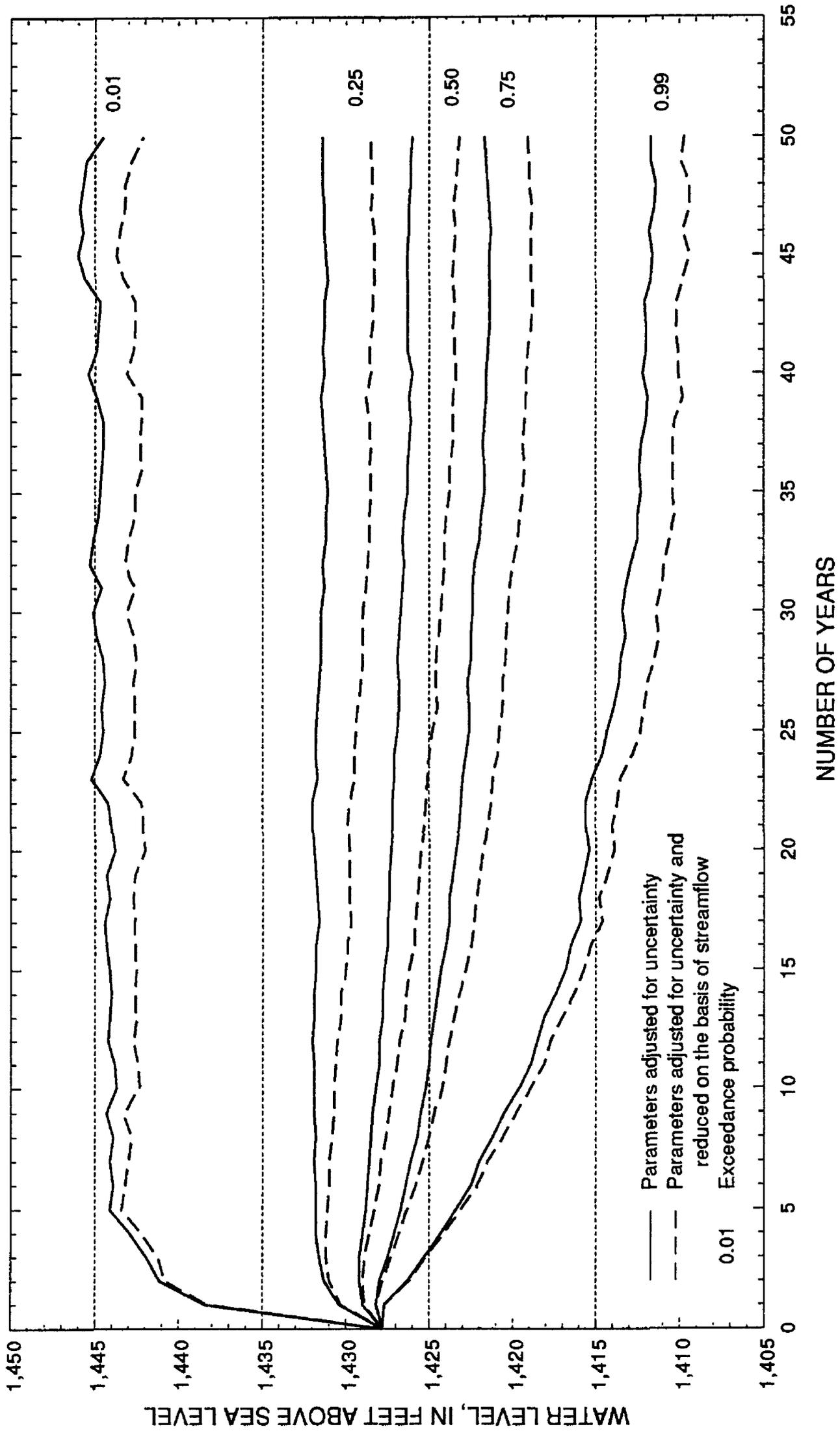


Figure 24. Exceedance levels for Devils Lake computed from 2,000 50-year traces from the water mass-balance model using parameters adjusted for uncertainty and reduced on the basis of streamflow for the Red River of the North at Grand Forks. (Year 0 corresponds to 1993.)

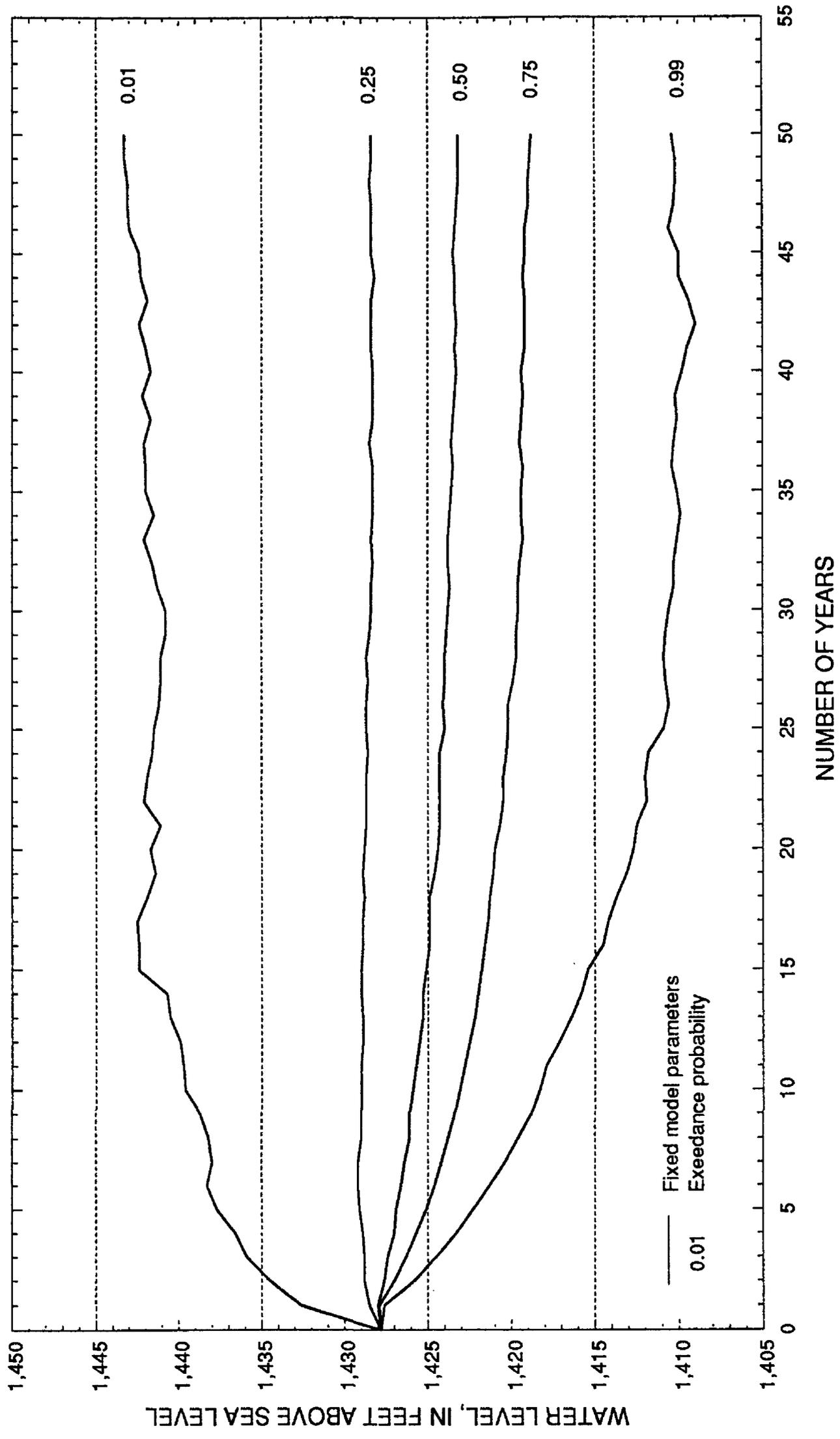


Figure 25. Exceedance levels for Devils Lake computed from 2,000 50-year traces from the water mass-balance model with each trace conditioned only on the starting lake level. (Year 0 corresponds to 1993.)

average (ARMA) time-series models and variable normality transformations were used to model both $X(t)$ and $D(t)$. Both models have a first-order autoregressive and first-order moving-average component. A square-root transformation was applied to $X(t)$ and a fourth-root transformation was used for $D(t)$. Residuals from the model for $X(t)$ were white noise (uncorrelated with zero mean and constant variance) and approximately normally distributed. Residuals from the model for $D(t)$ also were white noise but were not normally distributed. Nonnormality in the residuals for $D(t)$ was modeled by including a nonlinear relation between the residuals from the two models. Adjusted standardized residuals, $R^*(t)$, were close to a normal distribution.

Because of high autocorrelation (temporal persistence) in the lake levels of Devils Lake, the ALV model was verified using annual lake-level changes. The model closely reproduces the statistics of the recorded lake-level changes except for the skewness coefficient. However, the model output is less skewed than the data indicate because of some unrealistically large lake-level declines in the simulated data. For example, 1 percent of the lake-level changes for the simulated data were less than -3.3 feet (lake-level decline greater than 3.3 feet). The largest recorded lake-level decline for 1901-93 was -2.4 feet.

To estimate lake-level exceedance probabilities, 2,000 conditional simulation lake-level traces, each 50 years in length, were generated from the ALV model using parameters fixed at their fitted values. The high temporal persistence in the lake levels occurs because the effects of the initial conditions on the exceedance levels persist for several decades. The unconditional simulation exceedance levels are obtained in the limit as the length of the simulations increases. Therefore, unconditional simulation frequency analysis is of little use except in dealing with time scales of several centuries.

To account for estimation error in the parameters of the fitted ALV model, a parameter-uncertainty distribution was developed. Model parameters for each 50-year trace were sampled randomly from the parameter-uncertainty distribution rather than held fixed. The parameter uncertainty significantly affects the resulting exceedance levels. For example, the 1-percent exceedance level from the model that includes parameter uncertainty is 4.6 feet higher in year 50 than the 1-percent exceedance level from the model that has fixed parameters. The effect becomes even more pronounced at more extreme exceedance levels.

The statistical WMB model requires as inputs seasonal precipitation, evaporation, and inflow data for Devils Lake. Although the final WMB model was fitted on the basis of a quarterly time scale, the required inputs were computed using a monthly time scale. Precipitation data for the city of Devils Lake for 1950-93 were used in the WMB model. Lake evaporation for April through September was estimated by using monthly pan-evaporation data for Mandan and Dickinson. An apparent step-like increase in lake evaporation for Devils Lake occurred in 1959. Multiple linear regression was used to estimate the magnitude of the change in evaporation and to adjust the monthly pan-evaporation values for 1950-58. Monthly inflow to Devils Lake for 1950-93 was computed using a water mass-balance model. Monthly inflow was estimated on the basis of storage change, evaporation, lake-surface area, and precipitation.

Kendall's τ , Spearman's ρ , and Pearson's r were used to test for trends in the annual precipitation, evaporation, and inflow time series. The p values for Pearson's r for the annual precipitation and inflow time series are significant at the 10-percent level. However, the larger p values for Kendall's τ and Spearman's ρ indicate no upward trend in precipitation or inflow. The annual precipitation, evaporation, and inflow time series were assumed to be stationary.

The Davies and Harte test statistic (1987) was used to test for long-range dependence in the annual precipitation, evaporation, and inflow time series. No test statistics were significant at the 10-percent level. Thus, the assumption was made that no long-range dependence exists in the inputs.

Because the quarterly time scale was sufficient to generate realistic lake-level traces, the WMB model was based on quarterly rather than monthly data. Monthly precipitation, evaporation, and inflow data were aggregated to a quarterly (3-month) time scale and the model was fitted to the quarterly data using multiple linear regression. The model can be used to generate future lake-level traces given future precipitation, evaporation, and inflow time series. In the time-series model developed for quarterly precipitation, evaporation, and inflow, normality transformations were used to approximately maintain the marginal probability distributions; and a multivariate, periodic autoregressive model was used to reproduce the correlation structure. The model closely reproduces the statistics of the recorded seasonal inputs for 1950-93 and the statistics of the recorded lake-level data for Devils Lake. The unrealistically large lake-level declines produced using the ALV model were eliminated in the WMB model.

The WMB model was used to generate 2,000 conditional simulation lake-level traces, each 50 years in length. Model parameters for each trace were held fixed at their fitted values. A comparison of the exceedance levels from the WMB model and the exceedance levels from the ALV model indicates that the upper exceedance levels (25 percent and above) at the end of the simulation period agree fairly well but the lower levels do not.

The effect of parameter uncertainty in the fitted WMB model was included to determine the reliability of the exceedance levels. The parameters for each trace were sampled randomly from the parameter-uncertainty distribution rather than held fixed at their fitted values. A comparison of the 1-percent exceedance levels for the 2,000 50-year lake-level traces from the model that has fixed parameters and the 2,000 50-year lake-level traces from the model that includes parameter uncertainty indicates that the upper exceedance levels are particularly sensitive. The sensitivity in the upper exceedance levels is caused almost entirely by uncertainty in the fitted probability distributions of the quarterly inflows.

To reduce parameter uncertainty in the seasonal inflow distributions, long-term streamflow data for the Red River of the North at Grand Forks were used to reduce the variance in the estimated mean. The only significant effect of the adjustment was to reduce the estimated mean of the transformed spring inflow as well as the standard error of the estimated mean. The adjustment essentially lowered the stationary distribution of the lake levels by about 2 feet. For example, the 50-percent exceedance level for year 50 decreased from 1,425.7 to 1,423.0. The adjusted value of 1,423.0 is in agreement with the corresponding value from the ALV model, which adds confidence that both models are giving accurate estimates of the stationary exceedance levels, at least near the center of the distribution.

The WMB model required more extensive data collection and model development efforts than the simpler ALV model. However, there are several advantages to the WMB model that make it superior to the ALV model for estimating future lake-level probabilities of Devils Lake. The WMB model produces lake-level traces that closely follow the hydrology of Devils Lake as reflected by the water mass-balance equation. Because the lake-level traces from the ALV model are not constrained to follow the water mass-balance equation, many of the lake-level changes produced by the ALV model are unrealistic in terms of the amount of precipitation, evaporation, or inflow that would be required to produce those changes. The WMB model can be modified easily in future analyses to simulate the effects of various flood-control measures on the lake-level probabilities. The ALV model is not easily suited for simulating the effects of flood-control measures. Finally, the conditional simulation lake-level traces from the WMB model are much more sensitive to antecedent conditions than the conditional simulation lake-level traces from the ALV model. For example, the record precipitation and inflow in the summer and fall of 1993 caused the exceedance levels computed from the WMB model to increase much more rapidly in the first 20 years of the simulation period than the exceedance levels computed from the ALV model. The ALV model was not sensitive enough to predict the rapid rise because it is based only on annual lake volumes. Although the precipitation and inflow in the summer and fall of 1993 were extreme, the total annual precipitation and inflow for 1993 were not unusually large.

REFERENCES

- Aronow, Saul, 1955, Problems in late Pleistocene and Recent history of the Devils Lake region, North Dakota: Madison, University of Wisconsin, unpublished Ph.D. dissertation, 125 p.
- 1957, On the postglacial history of the Devils Lake region, North Dakota: *Journal of Geology*, v. 65, no. 4, p. 410-427.
- Bluemle, J.P., ed., 1981, The origin of Devils Lake: *North Dakota Geological Survey Newsletter*, December 1981, p. 7-9.
- 1988, Radiocarbon dates on Devils Lake beaches: *North Dakota Geological Survey Newsletter*, June 1988, p. 39-45.
- Bowles, D.S., and James, L.D., 1986, Issues associated with stochastic modeling of Great Salt Lake levels for planning purposes, *in* Kay, P.A., and Diaz, H.F., eds., *Problems of and prospects for predicting Great Salt Lake levels*: Center for Public Affairs Administration, University of Utah, May 1985, p. 218-235.
- Box, G.E.P., and Cox, D.R., 1964, An analysis of transformations: *Journal of the Royal Statistical Society*, v. B-26, p. 244-252.
- Box, G.E.P., and Jenkins, G.M., 1976, *Time series analysis; Forecasting and control*: San Francisco, Holden-Day, 575 p.
- Brockwell, P.J., and Davis, R.A., 1987, *Time series; Theory and methods*: New York, Springer-Verlag, 519 p.
- Brooks, C.E.P., 1951, Geological and historical aspects of climatic change, *in* Malone, T.F., ed., *Compendium of meteorology*: Boston, American Meteorological Society, p. 1004-1018.
- Callender, Edward, 1968, The postglacial sedimentology of Devils Lake, North Dakota: Grand Forks, University of North Dakota, unpublished Ph.D. dissertation, 312 p.
- Davies, R.B., and Harte, D.S., 1987, Tests for Hurst effect: *Biometrika*, v. 74, p. 95-101.
- Farnsworth, R.K., and Thompson, E.S., 1982, Mean monthly, seasonal, and annual pan evaporation for the United States: National Oceanic and Atmospheric Administration Technical Report NWS 33, 26 p.
- Helsel, D.R., and Hirsch, R.M., 1992, *Statistical methods in water resources*: Amsterdam, Elsevier, 522 p.
- Hurst, H.E., 1951, Long term storage capacity of reservoirs: *Transactions of the American Society of Civil Engineers*, v. 116, p. 770-799.
- James, L.D., Bowles, D.S., and Burges, S.J., 1977, Stochastic modeling for predicting frequency distributions of stages in terminal lakes at future planning horizons, *in* Greer, D.C., ed., *International Conference on Desertic Terminal Lakes*, Ogden, Utah, May 1977, Proceedings: Logan, Utah State University, Utah Water Resource Laboratory, p. 81-89.
- Mandelbrot, B.B., and Van Ness, J.W., 1968, Fractional Brownian motions, fractional noises and applications: *SIAM Review*, v. 10, p. 422-437.
- Mandelbrot, B.B., and Wallis, J.R., 1969, Robustness of the rescaled range R/S in the measurement of non-cyclic long run statistical dependence: *Water Resources Research*, v. 5, p. 967-988.
- Mitten, H.T., Scott, C.H., and Rosene, P.G., 1968, Chemical quality of surface waters in Devils Lake Basin, North Dakota, 1952-60: U.S. Geological Survey Water-Supply Paper 1859-B, 42 p.

- Mood, A.M., Graybill, F.A., and Boes, D.C., 1974, Introduction to the theory of statistics: New York, McGraw-Hill, 564 p.
- Pagano, M., 1978, On periodic and multiple autoregressions: *Annals of Statistics*, v. 6, p. 1310-1317.
- Pusc, S.W., 1993, Interaction between ground water and a large terminal lake, Devils Lake, North Dakota: Hydrogeology of Devils Lake area: North Dakota State Water Commission Water Resources Investigation 13, 95 p.
- Rowher, Carl, 1931, Evaporation from free water surfaces: U.S. Department of Agriculture Technical Bulletin Relating to Irrigation, No. 271.
- Ryan, G.L., and Wiche, G.J., 1988, Hydrology of the chain of lakes tributary to Devils Lake and water-level simulations of Devils Lake, northeastern North Dakota: U.S. Geological Survey Water-Resources Investigations Report 88-4020, 39 p.
- Salas, J.D., Tabois, G.Q., and Bartolini, P., 1985, Approaches to multivariate modeling of water resources time series: *Water Resources Bulletin*, v. 21, p. 683-708.
- StatSci, 1993, S-plus for windows user's manual: Statistical Sciences, Seattle, Washington.
- Stedinger, J.R., Pei, D., and Cohn, T.A., 1985, A condensed disaggregation model for incorporating parameter uncertainty into monthly reservoir simulations: *Water Resources Research*, v. 21, p. 665-675.
- Swenson, H.A., and Colby, B.R., 1955, Chemical quality of surface waters in Devils Lake Basin, North Dakota: U.S. Geological Survey Water-Supply Paper 1295, 82 p.
- U.S. Army Corps of Engineers, 1984, Reconnaissance report--Devils Lake Basin, North Dakota: St. Paul, Minnesota, 31 p.
- U.S. Department of Agriculture, Soil Conservation Service, [n.d.], Hydrology manual for North Dakota: Bismarck, North Dakota.
- U.S. Department of Agriculture, Weather Bureau, 1932a, Climatic summary of the United States, Western North Dakota: Washinton, D.C., sec. 34, 20 p.
- U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data Service, 1951-94, Climatological data, North Dakota, Annual summaries 1950-93: Asheville, North Carolina, v. 59-102, no. 13.
- 1955-64, Climatological data, North Dakota, Monthly summaries 1955-64: Asheville, North Carolina, v. 64-73, nos. 1-12.
- U.S. Water Resources Council, 1981, Guidelines for determining flood flow frequency: Washington, D.C., Hydrology Committee Bulletin 17B, 28 p.
- Upham, Warren, 1895, The glacial Lake Agassiz: U.S. Geological Survey Monograph No. 25, 658 p.
- Wiche, G.J., 1986, Hydrologic and climatologic factors affecting water levels of Devils Lake, North Dakota: U.S. Geological Survey Water-Resources Investigations Report 86-4320, 62 p.
- 1992, Evaporation computed by energy-budget and mass-transfer methods and water-balance estimates for Devils Lake, North Dakota, 1986-88: North Dakota State Water Commission Water Resources Investigation 11, 52 p.

- Wiche, G.J., Hoetzer, S.M., and Rankl, J.G., 1986, Hydrology of the Devils Lake Basin, northeastern North Dakota: North Dakota State Water Commission Water Resources Investigation 3, 86 p.
- Woodbury, L.H., and Padmanabhan, G., 1989, Estimating terminal lake level frequencies: *Journal of Water Resources Planning and Management*, v. 115, no. 3, May 1989, p. 321-336.
- Yevjevich, Vujica, 1982, Determination method for computing the probabilities of exceedance of future levels of Devils Lake, North Dakota: Contract DACW 37-82-M-1562, June 9, 1982.

SUPPLEMENT 1. PARAMETER-UNCERTAINTY DISTRIBUTIONS AND MODEL FITTING PROCEDURE

Parameter-Uncertainty Distribution for the Annual Lake-Volume Model

The annual lake-volume (ALV) model is based on autoregressive moving average (ARMA) models for the annual mean lake volume and the difference between the annual maximum lake volume and the annual mean lake volume (eqs. 7 and 9). Uncertainty in the fitted parameters of both models must be considered in order to assess the reliability of the lake-level frequency analysis on the basis of output of the ALV model. The model for the annual mean lake volume (eq. 7) can be expressed as

$$X'(t) - \mu = \phi [X'(t-1) - \mu] + \theta W(t-1) + W(t), \quad (\text{A1})$$

where

$$\begin{aligned} \hat{\mu} &\text{ is } 41.12, \\ \hat{\phi} &\text{ is } 0.967, \\ \hat{\theta} &\text{ is } 0.610, \text{ and} \\ \hat{\sigma}_w^2 &\text{ is } 4.624. \end{aligned}$$

Using the theory of maximum-likelihood estimation for fitting ARMA models (Brockwell and Davis, 1987, ch. 8), the estimated mean, $\hat{\mu}$, has an approximate normal distribution with the mean and variance given by

$$E(\hat{\mu}) = \mu \text{ and } Var(\hat{\mu}) = \frac{\sigma_w^2 (1 + \theta)^2}{n (1 - \phi)^2}. \quad (\text{A2})$$

The estimated autoregressive and moving-average parameters have an approximate bivariate, normal distribution with the mean and variance-covariance matrix given by

$$E[\hat{\phi}, \hat{\theta}] = [\phi, \theta] \text{ and } Var[\hat{\phi}, \hat{\theta}] = \frac{1}{n} \begin{bmatrix} (1 - \phi^2)^{-1} & (1 + \phi\theta)^{-1} \\ (1 + \phi\theta)^{-1} & (1 - \theta^2)^{-1} \end{bmatrix}^{-1}. \quad (\text{A3})$$

The estimated residual variance has an approximate normal distribution with the mean and variance given by

$$E(\hat{\sigma}_w^2) = \sigma_w^2 \text{ and } Var(\hat{\sigma}_w^2) = \frac{2}{n} \sigma_w^4. \quad (\text{A4})$$

Furthermore, the estimated mean, the estimated autoregressive and moving-average parameters, and the estimated residual variance are approximately independent of one another. These results are asymptotic (large sample) approximations that should be accurate considering the relatively large sample size ($n = 98$) used for parameter estimates.

To obtain the parameter-uncertainty distribution, the roles of the parameter estimates and the actual parameters are pivoted or reversed. The actual parameters are treated as random variables and the parameter estimates are treated as fixed quantities. The pivoting step is similar to the Bayesian approach used by Stedinger and others (1985) in which a posterior distribution of the model parameters is derived with respect to a noninformative prior distribution. Although the interpretation of the parameter-uncertainty distribution derived here using classical maximum-likelihood estimation results is different than the Bayesian interpretation used by Stedinger and others (1985), the parameter-uncertainty distributions obtained from the two approaches are similar.

The parameters in equations A2 to A4 were pivoted and the estimated parameter values from equation A1 were used to obtain the following parameter-uncertainty distributions for $X'(t)$:

$$\mu: \text{Normal}, E(\mu) = 41.12, \text{Var}(\mu) = (10.65)^2; \quad (\text{A5})$$

$$[\phi, \theta]: \text{Bivariate normal}, E(\phi) = 0.967, \text{Var}(\phi) = (0.026)^2, \quad (\text{A6})$$

$$E(\theta) = 0.610, \text{Var}(\theta) = (0.081)^2, \text{Corr}(\phi, \theta) = 0.127;$$

and

$$\sigma_w^2: \text{Normal}, E(\sigma_w^2) = 4.624, \text{Var}(\sigma_w^2) = (0.664)^2. \quad (\text{A7})$$

The three distributions are mutually independent. The largest parameter uncertainty is associated with the mean. For example, 95 percent of the values of μ generated from equation A5 are between 20.25 and 61.99. If equation 4 (with $\lambda = 1/2$) is used to convert these values to original units, 95 percent of the generated means of $X(t)$ are between 123,800 acre-feet and 1,024,000 acre-feet. This large range of uncertainty is caused by the high temporal persistence in the annual average lake-volume time series and causes the fitted autoregressive coefficient in equation A1 to be close to 1.0. Uncertainty in the fitted autoregressive coefficient also has a strong effect on the generated model output. The marginal distribution of ϕ in equation A6 is a normal distribution with a mean of 0.967 and a standard deviation of 0.026. Therefore, about 10 percent of the values of ϕ generated from equation A6 will exceed 1.0. Because ϕ must have an absolute value of less than 1.0 in order for the stationary distribution of $X'(t)$ to exist (Brockwell and Davis, 1987), parameter sets in which $\phi \geq 1.0$ were not used. This reflects the assumption that the annual average lake-volume time series is stationary.

The model for the difference between the annual maximum lake volume and the annual mean lake volume (eq. 9) can be expressed as

$$D'(t) - \mu = \phi [D'(t-1) - \mu] + \theta Z(t-1) + Z(t), \quad (\text{A8})$$

where

$$\begin{aligned} \hat{\mu} &\text{ is } 4.912, \\ \hat{\phi} &\text{ is } 0.943, \\ \hat{\theta} &\text{ is } -0.600, \text{ and} \\ \hat{\sigma}_Z^2 &\text{ is } 2.110. \end{aligned}$$

The parameters in equations A2 to A4 were pivoted and the fitted parameter values from equation A8 were used to obtain the following parameter-uncertainty distributions for $D'(t)$:

$$\mu: \text{Normal}, E(\mu) = 4.912, \text{Var}(\mu) = (1.057)^2; \quad (\text{A9})$$

$$[\phi, \theta]: \text{Bivariate normal}, E(\phi) = 0.943, \text{Var}(\phi) = (0.044)^2, \quad (\text{A10})$$

$$E(\theta) = -0.600, \text{Var}(\theta) = (0.105)^2, \text{Corr}(\phi, \theta) = -0.608;$$

and

$$\sigma_z^2: \text{Normal}, E(\sigma_z^2) = 2.110, \text{Var}(\sigma_z^2) = (0.309)^2. \quad (\text{A11})$$

As with $X'(t)$, about 10 percent of the values of ϕ generated from equation A10 will exceed 1.0. Parameter sets in which this occurs are not used in order to ensure that $D'(t)$ is stationary.

To obtain a lake-level trace from the ALV model that includes parameter uncertainty, parameters from the ARMA models of $X(t)$ and $D(t)$ are generated from the previous parameter-uncertainty distributions. A new set of parameters is generated for each new trace.

Model Fitting Procedure for Quarterly Precipitation, Evaporation, and Inflow

The seasonal Yule-Walker equations (Salas and others, 1985) were used to obtain estimates of the autoregressive parameters in the multivariate, periodic autoregressive model for the standardized quarterly precipitation, evaporation, and inflow (eqs. 35 to 46). In order to define the estimates, the quarterly data for all three inputs were combined into a single, one-dimensional array as follows:

$$W = [P^*(1), E^*(1), Q^*(1); P^*(2), E^*(2), Q^*(2); \dots; P^*(176), E^*(176), Q^*(176)]. \quad (\text{A12})$$

Each input time series has 176 observations (44 years of data times 4 quarters per year) and W has 528 (176x3) total elements. $W(j)$ is defined as the j th element of W . For example, $W(6)$ corresponds to spring inflow for 1950, $W(18)$ corresponds to spring inflow for 1951, and so on. To define the estimates, each quarter and each input variable were considered separately. For example, spring inflow, which corresponds to elements $\{W(6 + 12k), k = 0, 1, \dots, 43\}$, can be expressed as

$$W(6 + 12k) = \sum_{j=1}^p \hat{\phi}_j W(6 + 12k - l_j) + \hat{\sigma} Z(6 + 12k); k = 0, 1, \dots, 43, \quad (\text{A13})$$

where

$$[\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p] = r_p^T R_p^{-1}; \hat{\sigma}^2 = 1 - r_p^T R_p^{-1} r_p \text{ and} \quad (\text{A14})$$

l_j is the j th lag.

R_p is a ($p \times p$) matrix with ij th element as follows:

$$[R_p]_{ij} = \frac{1}{44} \sum_{k=0}^{43} W(6 + 12k - l_i) W(6 + 12k - l_j). \quad (\text{A15})$$

r_p is a ($p \times 1$) vector with i th element as follows:

$$[r_p]_i = \frac{1}{44} \sum_{k=0}^{43} W(6 + 12k) W(6 + 12k - l_i). \quad (\text{A16})$$

In equations A13 to A16, $W(j)$ should be set to zero whenever $j < 1$. The method for determining the individual lagged inputs, l_j , for the model will be given later. Pagano (1978) showed that the seasonal Yule-Walker estimators defined in equation A14 have an approximate normal distribution with the mean and covariance matrix given by

$$E[\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p] = [\phi_1, \phi_2, \dots, \phi_p] \text{ and} \quad (\text{A17})$$

$$\text{Cov}[\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p] = \frac{\hat{\sigma}^2}{n} R_p^{-1}.$$

A parameter estimate, $\hat{\phi}_j$, is significant at the 0.05 level if $|\hat{\phi}_j| / \sqrt{\text{Var}(\hat{\phi}_j)} > 1.96$ and the variance of the estimate is obtained from the j th diagonal element of the covariance matrix in equation A17.

Model variables were selected by a stepwise variable inclusion procedure. The procedure is demonstrated for spring inflow as follows:

1. Set $p = 1$ and fit the model (eq. A13) for each value of $l_1 = 1, 2, \dots, 12$. A maximum value of 12 was selected to allow lags of up to 1 year in the model.
2. Select the value of l_1 that results in the smallest residual variance. If the corresponding autoregressive coefficient is significant at the 0.05 level, keep the coefficient in the model and continue to step 3. Otherwise, set $p = 0$ and stop.
3. Given that $W(6 + 12k - l_1)$ is in the model, set $p = 2$ and fit the model (eq. A13) for $l_2 = 1, 2, \dots, 12$, excluding the value obtained for l_1 .
4. Select the value of l_2 that results in the smallest residual variance. If the corresponding autoregressive coefficient is significant at the 0.05 level, keep the coefficient in the model and continue to step 5. Otherwise, set $p = 1$ and stop.
5. Continue in like manner until all significant autoregressive coefficients are included in the model.

The variable inclusion procedure for spring inflow resulted in $p = 3$, $l_1 = 5$, $l_2 = 6$, and $l_3 = 12$. The fitted model was “decoded” by converting the elements of W back to the individual lagged inputs. The resulting model is given by equation 40. The same procedure was applied to each of the seasons and each of the variables, resulting in the final fitted model (eqs. 35 to 46).

Parameter-Uncertainty Distribution for the Water Mass-Balance Model

The parameter-uncertainty analysis for the water mass-balance (WMB) model is complicated by the relatively large number of parameters in the final fitted model. The parameter estimates in the uncertainty analysis include 12 sample means and 12 sample variances of the Box-Cox transformed inputs (eqs. 23 to 34) and 16 autoregressive parameters in the multivariate, periodic autoregressive model (eqs. 35 to 46).

Uncertainty in the sample means will be considered first. To simplify the derivation, a vector-valued annual time series from the elements given in equation A12 is defined as

$$U_k = \{W(1 + 12k), W(2 + 12k), \dots, W(12 + 12k)\}^T, \quad k = 0, 1, \dots, 43. \quad (\text{A18})$$

A similar array from the elements of the unstandardized Box-Cox transformed inputs is defined as

$$W' = [P'(1), E'(1), Q'(1); P'(2), E'(2), Q'(2); \dots; P'(176), E'(176), Q'(176)]. \quad (\text{A19})$$

Let

$$U'_k = [W'(1 + 12k), W'(2 + 12k), \dots, W'(12 + 12k)]^T, \quad k = 0, 1, \dots, 43. \quad (\text{A20})$$

The sample means in equations 23 to 34 can be expressed in vector form as

$$\hat{\mu} = \frac{1}{44} \sum_{k=0}^{43} U'_k = [1.19, 1.43, \dots, 9.57]^T; \quad (\text{A21})$$

and the multivariate, periodic autoregressive model given by equations 35 to 46 can be expressed in vector form as

$$LU_k = AU_{k-1} + DZ_k, \quad (\text{A22})$$

where

L and A are (12x12) matrices whose elements depend on the autoregressive parameters, D is a diagonal matrix with the residual standard deviations on the diagonal; and Z_k is a vector of independent, standard normal, random variables.

The particular form of L , A , and D can be determined from equations 35 to 46. L and A are both sparse matrices that have many zero elements. Using equation A21, equation A22, and results from Brockwell and Davis (1987, ch. 11), the vector of sample means has a normal distribution with the mean and approximate covariance matrix given by

$$E(\hat{\mu}) = \mu, \quad Cov(\hat{\mu}) = \frac{1}{n} S(L - A)^{-1} D^2 [(L - A)^{-1}]^T S, \quad (\text{A23})$$

where

L , A , and D are defined in equation A22;
 n is the number of years of data (44); and
 S is a diagonal matrix,

$$S = \left(\text{diag} \left[\frac{1}{44} \sum_{k=0}^{43} (U'_k - \hat{\mu}) (U'_k - \hat{\mu})^T \right] \right)^{1/2} = \text{diag} [0.19, 0.24, \dots, 0.47]; \quad (\text{A24})$$

that has the estimated standard deviations of the Box-Cox transformed inputs on the diagonal. The square roots of the diagonal elements of the covariance matrix in equation A23 are the estimated standard deviations of the sample means of the Box-Cox transformed inputs. These values are given in table A1. The estimated crosscorrelation matrix of the sample means is obtained by converting the covariance matrix in equation A23 into a crosscorrelation matrix. The resulting crosscorrelation matrix is given in table A2.

Table A1. Sample means and estimated standard deviations of the sample means for Box-Cox transformed quarterly precipitation, evaporation, and inflow

Quarter	Variable	Sample mean	Estimated standard deviation of sample mean
Winter	P	1.19	0.029
	E	1.43	.036
	Q	¹ 6,051.50	769.242
Spring	P	1.88	.030
	E	1.57	.036
	Q	7.03	.468
Summer	P	1.85	.038
	E	18.02	.282
	Q	9.81	.148
Fall	P	1.27	.038
	E	4.22	.055
	Q	9.57	.081

¹Value was not transformed.

To simulate uncertainty in the sample means, the vector of 12 means for each trace is generated from a multivariate, normal distribution. Expected values for the sample means and estimated standard deviations are given in table A1. The estimated crosscorrelation matrix is given in table A2.

Uncertainty in the fitted autoregressive parameters in equations 35 to 46 will be considered next. The autoregressive parameters from any particular row of the model have an approximate multivariate, normal distribution with a mean and covariance matrix as given in equation A17. The estimated standard deviations and crosscorrelation matrices are given in table A3. Pagano (1978) showed that the set of autoregressive parameters from any particular row of the model are asymptotically independent of the autoregressive parameters from any other row.

Table A2. Estimated crosscorrelation matrix of sample means for Box-Cox transformed quarterly precipitation, evaporation, and inflow

<i>P1</i>	0	0	0	0	0.48	0	0	0.29	0	0	0.15
	<i>E1</i>	0	0	.50	.04	0	.49	.21	0	.62	.11
		<i>Q1</i>	0	0	.05	0	0	.03	.46	0	.15
			<i>P2</i>	.02	0	0	0	0	0	.01	0
				<i>E2</i>	.02	0	.25	.11	0	.69	.06
					<i>Q2</i>	.02	.05	.71	.11	.05	.60
						<i>P3</i>	-.46	.12	0	-.34	.06
							<i>E3</i>	.28	0	.86	.15
								<i>Q3</i>	.07	.26	.67
									<i>P4</i>	0	.32
										<i>E4</i>	.14
											<i>Q4</i>

Table A3. Estimated standard deviations and crosscorrelation matrices of autoregressive parameters in the fitted multivariate, periodic autoregressive model for quarterly precipitation, evaporation, and inflow

Model equation number	Fitted coefficient	Estimated standard deviation	Crosscorrelation matrix		
37	0.47	0.13	(¹)	(¹)	(¹)
39	.48	.13	1.00	0.19	0.02
	.31	.13	.19	1.00	-.06
	-.30	.12	.02	-.06	1.00
40	.46	.12	1.00	.06	-.20
	.32	.12	.06	1.00	-.21
	.21	.11	-.20	-.21	1.00
42	.49	.11	1.00	-.04	(²)
	-.46	.11	-.04	1.00	(²)
43	.54	.11	1.00	.11	.20
	.46	.12	.11	1.00	.46
	.34	.13	.20	.46	1.00
45	.73	.02	1.00	.21	(²)
	.52	.02	.21	1.00	(²)
46	.50	.12	1.00	.14	(²)
	.32	.12	.14	1.00	(²)

¹One-by-one matrix.

²Two-by-two matrix.

To simulate uncertainty in the autoregressive parameters, the parameters of each trace are generated from a multivariate, normal distribution. The estimated standard deviations and crosscorrelation matrices obtained are given in table A3. The parameters of equations 35 to 46 are generated independently.

Uncertainty in the sample variances of the Box-Cox transformed inputs is considered last. Two approaches can be used to include the uncertainty. In the first approach, the variances of the transformed inputs in equations 23 to 34 vary and the variances of each of the standardized inputs in equations 35 to 46 are equal to one. In this approach, the residual variances in equations 35 to 46 are fixed functions of the autoregressive parameters because of the constraint on the variances of the standardized inputs. In the second approach, the variances of the transformed inputs in equations 23 to 34 are fixed and the residual variances in equations 35 to 46 vary according to the proper parameter-uncertainty distribution. The second approach will be used in this report. Pagano (1978) showed that the estimated residual variance from any particular row of the model (eq. A14) has an approximate normal distribution with a mean and variance of

$$E(\hat{\sigma}^2) = \sigma^2 \text{ and } Var(\hat{\sigma}^2) = \frac{2}{n} \hat{\sigma}^4. \quad (\text{A25})$$

Furthermore, the estimated residual variances are independent of one another and of the estimated autoregressive parameters. The residual variances of the fitted model and the estimated standard deviations are given in table A4. To simulate uncertainty in the residual variances, the residual variances of each trace are sampled from an independent, normal distribution.

Table A4. Residual variances and estimated standard deviations of residual variances for the multivariate, periodic autoregressive model for quarterly precipitation, evaporation, and inflow

Model equation number	Residual variance	Estimated standard deviation of residual variance
35	1.00	0.21
36	1.00	.21
37	.78	.17
38	1.00	.21
39	.68	.14
40	.61	.13
41	1.00	.21
42	.54	.12
43	.53	.11
44	1.00	.21
45	.02	.004
46	.64	.14

Variance of Adjusted Means of Seasonal Inflow Distributions

The variance of the adjusted mean of the seasonal inflows (eq. 56) is derived from the following bivariate time-series model for log-transformed streamflow data for the Red River of the North at Grand Forks and Box-Cox transformed inflow data for Devils Lake for a particular season:

$$X(t) = (1 - \phi_x) \mu_x + \phi_x X(t-1) + \frac{\sigma_x}{\sqrt{1 - \phi_x^2}} \varepsilon(t) \quad (\text{A26})$$

$$Y(t) = (1 - \phi_y)\mu_y + \phi_y Y(t-1) + \frac{\sigma_y}{\sqrt{1 - \phi_y^2} \sqrt{1 - p^2}} [\eta(t) + p\varepsilon(t)]$$

where

$X(t)$ is log-transformed streamflow data for the Red River of the North at Grand Forks;

$Y(t)$ is Box-Cox transformed inflow data for Devils Lake;

$\phi_x = \text{Corr}[X(t), X(t-1)]$; $\mu_x = E[X(t)]$;

$\phi_y = \text{Corr}[Y(t), Y(t-1)]$; $\mu_y = E[Y(t)]$;

$p = \text{Corr}[X(t) - \phi_x X(t-1), Y(t) - \phi_y Y(t-1)]$; and

$\varepsilon(t)$ and $\eta(t)$ are independent, white noise time series with a mean of zero and a variance of 1.

Equation A26 is based on the assumption that $X(t)$ and $Y(t)$ each follow a first-order autoregressive model with cross-correlated residuals. The model (eq. A26) gives a good approximation of the recorded data from 1950-93.

The following expression was used to determine the constant, a , in equation 52 that minimizes the variance of the estimated mean:

$$\text{Var}(\hat{\mu}') = \text{Var}(\hat{\mu}) + a^2 \text{Var}(\bar{X}_1 - \bar{X}_2) + 2a \text{Cov}(\hat{\mu}, \bar{X}_1 - \bar{X}_2).$$

Minimizing this expression with respect to a , the following value of a is obtained along with the corresponding minimized variance:

$$a_m = \frac{-\text{Cov}(\hat{\mu}, \bar{X}_1 - \bar{X}_2)}{\text{Var}(\bar{X}_1 - \bar{X}_2)} = \frac{\text{Cov}(\hat{\mu}, \bar{X}_2) - \text{Cov}(\hat{\mu}, \bar{X}_1)}{\text{Var}(\bar{X}_1) + \text{Var}(\bar{X}_2) - 2\text{Cov}(\bar{X}_1, \bar{X}_2)} \quad (\text{A27})$$

$$\text{Var}[\hat{\mu} + a_m(\bar{X}_1 - \bar{X}_2)] = \text{Var}(\hat{\mu}) - a_m^2 \text{Var}(\bar{X}_1 - \bar{X}_2). \quad (\text{A28})$$

Assuming that the model (eq. A26) is correct and using large-sample approximations such as discussed in Brockwell and Davis (1987, p. 212), the following expressions are obtained:

$$\text{Var}(\hat{\mu}) = \frac{\sigma_y^2}{n_2} \frac{1 + \phi_y}{1 - \phi_y} + o\left(\frac{1}{n_2}\right)$$

$$\text{Var}(\bar{X}_i) = \frac{\sigma_x^2}{n_i} \frac{1 + \phi_x}{1 - \phi_x} + o\left(\frac{1}{n_i}\right)$$

$$\text{Cov}(\hat{\mu}, \bar{X}_1) = \text{Cov}(\bar{X}_2, \bar{X}_1) = o\left(\frac{1}{n_1}\right)$$

$$\text{Cov}(\hat{\mu}, \bar{X}_2) = \frac{p\sigma_x\sigma_y\sqrt{(1+\phi_x)(1+\phi_y)}}{n_2\sqrt{(1-\phi_x)(1-\phi_y)}} + o\left(\frac{1}{n_2}\right).$$

The approximations given in equations 54 and 56 are obtained by substituting these expressions into equations A27 and A28 and ignoring terms of smaller order than $1/n_2$.