DOCUMENT RESUME

ED 243 489 IR 050 715

AUTHOR	Brooks, Terrence A.; Forys, John W., Jr.
TITLE	Predicting Academic Library Circulations: A Forecasting Methods Competition.
SPONS AGENCY	Council on Library Resources, Inc., Washington, D.C.
PUB DATE	1 May 84
NOTE	115p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Academic Libraries; Goodness of Fit; Higher Education; *Library Circulation; Mathematical Models; *Prediction; *Predictive Measurement; *Statistical Data
IDENTIFIERS	*Library Statistics; Linear Trends; *Smoothing Methods

ABSTRACT

Based on sample data representing five years of monthly circulation totals from 50 academic libraries in Illinois, Iowa, Michigan, Minnesota, Missouri, and Ohio, a study was conducted to determine the most efficient smoothing forecasting methods for academic libraries. Smoothing forecasting methods were chosen because they have been characterized as easy to use and fairly accurate. It was found that smoothing forecasting methods worked very poorly on monthly library data due to the seasonality present in monthly library circulation totals. The only method recommended for use with monthly data was Winters' Linear and Seasonal Exponential Smoothing method, which has a specific seasonal component. Much greater success was achieved by using smoothing forecasting methods with yearly-lagged data, for example, using the circulation totals of past Januarys to predict the total of a future January. The One-Month Single Moving Average was found to be the most efficient smoothing method for forecasting future monthly circulation totals on yearly-lagged data with litle or no trend, while Brown's One-Parameter Linear Exponential method (with alpha set at 0.5) was recommended for use in trending yearly-lagged data. These methods ranked first and second respectively in minimizing both the mean percentage forecasting error and standard deviation of forecasting errors. A 27-item bibliography and plots showing the circulation data from the 50 libraries are included. (ESR)

PREDICTING ACADEMIC LIBRARY CIRCULATIONS:

A FORECASTING METHODS COMPETITION

by

Terrence A. Brooks
School of Library and Information Science
University of Iowa
Iowa City, IA 52242

John W. Forys, Jr.
Engineering Library
University of Iowa
Iowa City, IA 52242

A Research Report Submitted to the
Council on Library Resources
1785 Massachusetts Avenue, N.W.
Washington, DC 20036

May 1, 1984

EXECUTIVE SUMMARY

This study identified forecasting methods as missing from the management skills of academic librarians. Smoothing forecasting methods were proposed as potential remedies. Forecasting researchers have characterized these techniques as easy to use and as fairly accurate for the time and effort invested in their calculation. They seem to be good candidate forecasting techniques for academic librarians to try. A forecasting competition was used to determine the most efficient smoothing forecasting methods for academic libraries. The sample data were five years of monthly circulation totals from fifty academic libraries. Plots of each library's data showed that the vast majority of libraries have heavily seasonal data.

The results showed that smoothing forecasting methods work very poorly on monthly library data and we recommend that librarians not use them on monthly library data. The reason for the poor performance of smoothing methods is the seasonality present in monthly library circulation totals. If a librarian does wish, however, to employ smoothing forecasting methods then we urge him to use Winters' Linear and Seasonal Exponential Smoothing which has a specific seasonal component.

Much greater success was achieved by using smoothing forecasting methods with yearly-lagged data. An example of

lagging data one year is using the circulation totals of past Januarys to predict the total of a future January. It appears that for many libraries taking a yearly-lagged approach completely de-seasonalizes monthly circulation data. For some libraries only a trend component is left in the data.

We recommended that for those libraries with little or no trend in their yearly-lagged data, the One-Month Single Moving Average be used for forecasting future monthly circulation totals. This method was the most efficient smoothing method in minimizing both the mean percentage error and standard deviation of forecasting errors. For trending yearly-lagged data, we recommended that academic librarians use Brown's One-Parameter Linear Exponential method (with alpha set at 0.5) for forecasting future monthly circulation totals. This method can adjust its forecasts when a trend is present. It ranked second in minimizing both the mean percentage forecasting error and standard deviation of forecasting errors.

ACKNOWLEDGEMENTS

Many people have contributed to the success of this endeavor. The largest group of unnamed and unheralded supporters are the librarians who took time from their busy schedules to meet our demands for data about their libraries. Those people have our many thanks.

The Council on Library Resources provided our financial support and thus in a very real way made this study possible. Both Dale Bentz, Librarian of the University of Iowa, and Carl Orgren, Director of the School of Library and Information Science, provided us with support and encouragement.

This report could not have been written without the daily efforts of Judy Butcher who spent long hours before a computer terminal.

Finally our gratitude goes to our families who generously forgave our enthusiasm for library statistics through these many months.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

There has been a renaissance of interest in quantitative forecasting in science, management, and economics with only limited transfer of these new forecasting techniques to the management of academic libraries. Forecasting is such a well-developed tool of business that advice has been offered to managers about which forecasting technique to choose [Chambers, Mullick, & Smith, 1971]. Unfortunately, academic librarians don't enjoy equivalent sage advice about which forecasting method to use. This study proposed to fill this gap in the knowledge base of academic librarians by applying forecasting techniques of the smoothing variety [Makridakis & Wheelwright, 1978, chapter 3] to a large sample of academic libraries' monthly circulations in a forecasting accuracy competition. The goal of this research study was to find the smoothing forecasting methods that are the most efficient forecasters of academic library circulation data.

Library literature reveals little awareness of forecasting. Instead, there are many statements made about the absence of quantitative tools in library decision making. Many authors lament the current state of library statistics. They have been characterized as primitive by Rogers and Weber [1971, p.275], and as busywork by Herner [1967, p.47]. Burns [1974] has characterized the collection of library statistics as a routine operation with unclear purpose that produces data that are utilized at a very unsophisticated level. However, Burns anticipated a better future: "There will be a more sophisticated use of statistics to measure, forecast, simulate, and model all phases of library operations especially those of circulation" [1977, p.7].

BACKGROUND ON FORECASTING

Hamburg [1978, p.38] stated that "in library planning and decision-making predictions are invariably required". His bold statement has not motivated much theoretical work or practical application of forecasting methodologies to library statistics. Even as the economic environment of libraries has worsened, library managers have not turned to techniques such as forecasting that would serve to fine tune a library's activities. "In particular, as resources dwindle libraries want to predict the behaviors of their users and potential users so that they may both plan and promote their activities" [Dervin, 1977].

The use of forecasting in library administration practice is in sharp contrast to the acceptance of forecasting in other disciplines. Forecasting, or trend analysis, is considered as an integral part of scientific management and rational decision making. Makridakis and Wheelwright describe forecasting as a tool that permits management to shield an organization from the vagaries of chance events and become more methodical in dealing with its environment [1978, p.5]. Like bureaucracies everywhere, academic libraries need tools that will enhance planning and rational decision making. One tool to help accomplish these managerial tasks is forecasting.

Forecasting should be of interest to librarians and information scientists for at least two reasons. The first reason stems from the argument about the desirability of managerial rationality. Like all managers, the library manager must allocate his scarce resources prudently, and make "his decisions based on his predictions of the effects of allocating varying amounts of resources to the different functions in the library" [House, 1974]. Prediction methods can potentially become one of the daily tools of the library manager.

The second reason is more theoretical. Library-output statistics such as circulation data are intrinsically interesting variables that merit their own investigation. Previous work [Brooks, 1981] has demonstrated that library-output statistics have surprising characteristics that are unanticipated by the folklore about them. Forecasting studies are only one methodology for studying library-output statistics. By demonstrating an ability to forecast library-output statistics, we prove that we understand some of the dynamics that are driving them.

Library literature is not distinguished by sophisticated applications of forecasting techniques. Many authors [Carpenter & Vasu, 1978; Hoadley & Clark, 1972; Simpson, 1975] writing about quantitative or statistical methods in librarianship ignore forecasting completely. The topic is not treated in Lancaster's The Measurement and Evaluation of Library Services and in another volume, Investigative Methods in Library and Information Science: An Introduction, Martyn and Lancaster cover only the Delphi technique which is a method for divining a consensus of opinion. Conspicuously absent are any inferential statistical forecasting techniques. Rutherford Rogers and David Weber described the managerial use of library statistics as primitive and then proceeded to prove it by discussing forecasting only in terms of the descriptive method of plotting trends on charts [1971, p.279]. They neglected to discuss any inferential technique that could establish the statistical significance of a graphic trend line. Stueart and Eastlick [1981], who treated forecasting in three paragraphs, also recommended only graphical methods.

There are, however, two other library forecasting studies of special note. Miriam Drake [1976] considered linear regression as a predictive technique. She concluded that straight trend lines are not the most efficient predictors in all library situations. The reason is that library data, especially circulation data, show monthly or seasonal fluctuations. Cyclicity may be one of the reasons that forecasting techniques have had a retarded application to library statistics. Cyclicity of monthly library totals certainly played a large part in this study. The reader is invited to peruse the many monthly circulation plots given in Appendix 3. Most of these data are strongly cyclical in that patterns that appear in one year are often repeated in other years.

The most sophisticated forecasting study in library literature to date is by Kang [1979]. He forecasted the requests for interlibrary loan services received by the Illinois Research and Reference Centers from 1971 through 1978 using several methods, including methods that can model cyclical data, and found regression to be the best predictive technique. He used a weighted regression formula that gave less predictive value to older observations, and greater weight to the most recent ones. His study is flawed by the fact that only one set of data was used; hence, the generalizability of Kang study is severely limited.

RELATED FORECASTING STUDIES

There are a number of related studies that have attempted to predict circulation with causal techniques. Hodowanec [1980] used multiple regression analysis with twenty independent variables to predict circulation patterns of graduate students, undergraduate students, and faculty. McGrath [1976-1977] isolated twelve independent variables in a multiple regression analysis to predict circulation of monographs by academic subjects. Zweizig [1973] used the related approach of multiple discriminant analysis to isolate factors that determine public library use.

Another approach to predicting future library use is modeling standard statistical distributions on a sample of library circulations. For example, Lazorick [1970] found the demand for books in a collection to follow a negative binomial distribution. Nozik [1974] used a Markov process to model book demand, and Burrell [1980] offered a model to show likely patterns of future use of individual book titles. Morse and Chen [1975] showed how bias can be controlled in predicting the total yearly circulation of each class of books in a library, and Slote [1970] studied the past use of individual books as a predictor of their future use. Other investigators have used random samples to predict the total number of patrons entering a library [Jones, 1973].

It is clear from this examination of related library literature that no study has ever used smoothing forecasting methods on academic library circulations.

DATA COLLECTION

Beginning in December 1982 and continuing in January 1983, fifty academic libraries from the American Library Directory were chosen randomly from the midwest states of Illinois, Iowa, Michigan, Minnesota, Missouri, and Ohio. A copy of the letter sent to each library is given in Appendix 1.

The data requested were total monthly circulation counts for a five-year period, i.e., 60 consecutive monthly circulation totals. The aim was to collect a set of time-series data from each library. Usable data were received from two libraries in Iowa, six libraries in Illinois, eight libraries in Michigan, one library in Minnesota, nine libraries in Missouri, and twenty-four libraries in Ohio.

These data were loaded in a computer and a plot was drawn for each library. The monthly circulation plots appear in Appendix 3. The reader is invited to consult the plotted data. It is our observation from these plots that in most instances academic libraries have strongly seasonal patterns in their monthly circulation totals. That is, it is evident in most of the plots that certain characteristic patterns repeat themselves from year to year.

SMOOTHING FORECASTING METHODOLOGY

The study applies the smoothing forecasting methods of Makridakis and Wheelwright [1978] to a sample of academic library circulation statistics. These authors present formulas for a number of smoothing methodologies.

A smoothing forecasting method uses the information supplied by previous data to create a forecast for the future. It is assumed in these methods that a signal exists in the past data that may be obscured by a certain small percentage of random errors. A smoothing forecasting method weights certain past observations and averages past random errors in order to reveal the underlying signal in the data. There are many methods and many ways to utilize the information of past observations. It is often the case that each method permits many variations in either the selection of weights or in the time lags used for smoothing. We have capitalized on this flexibility by systematically varying weights and time lags.

Because of the highly seasonal nature of monthly library statistics, we employed the following strategy: all the methods were run on the monthly data, and then the methods were run on data lagged one year. These yearly-lagged data were developed from the 60 data points available to us from each library and consisted of data points in positions 1, 13, 25, 37, and 49. These data points were treated as a separate, and smaller, time series from each library. Not every library donated data that began with January, thus the lagged time series don't represent series that jump from January to January to January, etc. Using data lagged in this fashion de-seasonalizes a monthly time-series because it transforms the data from a series of months that run from January to December throughout one year to a series of the identical months across several years. Thus it doesn't matter, for example, that July may be a quieter month than August. With yearly-lagged data, Julys are compared to each other, as are Augusts compared to each other, etc.

Following is a listing of the methods used in this study and the techniques used to initialize each method.

<u>Stylistic note:</u> In the following formulas braces have been used to denote subscripts. Thus F{t} should be read as F subscript t.

A. Single Moving Averages

Eight variations of single moving averages were used. An exposition of the single moving average technique is given by Wheelwright and Makridakis [1978, p.45].

1. One-Month Moving Average

$F\{t+1\}=X\{t\}$

Initialization:

$F\{2\}=X\{1\}$

Note: Last month's total is used as next month's forecast.

2. Two-Month Moving Average

$F\{t+1\}=[X\{t\}+X\{t-1\}]/2$

Initialization:

$F\{3\}=[X\{1\}+X\{2\}]/2$

Note: The average of the two preceding months is used to forecast.

3. Three-Month Moving Average

$F\{t+1\}=[X\{t\}+X\{t-1\}+X\{t-2\}]/3$

Initialization:

$F\{4\}=[X\{1\}+X\{2\}+X\{3\}]/3$

Note: The average of the three preceding months is used to forecast.

4. Four-Month Moving Average

F{t+1}=[X{t}+X{t-1}+X{t-2}+X{t-3}]/4

Initialization:

F{5}=[X{4}+X{3}+X{2}+X{1}]/4

Note: The average of the four preceding months is used to forecast.

5. One/12-Month Moving Average

F{t+12}=X{t}

Initialization:

F{13}=X{1}

Note: Permitting a month's total this year to be the forecast for the same month next year removes the effect of seasonality in the data.

6. Two/12-Month Moving Average

F{t+12}=[X{t}+X{t-12}]/2

Initialization:

F{25}=[X{13}+X{1}]/2

Note: The first and thirteenth observations are averaged to provide a forecast for the twenty-fifth observation.

7. Three/12-Month Moving Average

$F\{t+12\}=[X\{t\}+X\{t-12\}+X\{t-24\}]/3$

Initialization:

$F\{37\}=[X\{25\}+X\{13\}+X\{1\}]/3$

Note: Observations from three preceeding years are averaged to provide a forecast for the fourth year.

8. Four/12-Month Moving Average

$F\{t+12\}=[X\{t\}+X\{t-12\}+X\{t-24\}+X\{t-36\}]/4$

Initialization:

$F\{49\}=[X\{37\}+X\{25\}+X\{13\}+X\{1\}]/4$

Note: Observations from four preceeding years are averaged to provide a forecast for the fifth year.

B. Linear Moving Averages

Four variations of the linear moving average method were used. An exposition of the linear moving average technique is given by Wheelwright and Makridakis [1978, p.55]

1. Two-Month Linear Moving Average-Lag One Month

$F\{t+1\}=a\{t\}+b\{t\}m$

where $m = 1$ and:

$b\{t\}=2(S'\{t\}-S''\{t\})$

$a\{t\}=2S'\{t\}-S''\{t\}$

$S''\{t\}=[S'\{t\}+S'\{t-1\}]/2$

$S'\{t\}=[X\{t\}+X\{t-1\}]/2$

Initialization:

$S'\{2\}=[X\{2\}+X\{1\}]/2$

$S'\{3\}=[X\{3\}+X\{2\}]/2$

$S''\{3\}=[S'\{3\}+S'\{2\}]/2$

$a\{3\}=2S'\{3\}-S''\{3\}$

$b\{3\}=2(S'\{3\}-S''\{3\})$

$F\{4\}=a\{3\}+b\{3\}$

Note: In this method two observations are averaged or smoothed. This average itself is then averaged with a previously calculated average to create a double-smoothed effect.

2. Three-Month Linear Moving Average-Lag One Month

$F\{t+1\}=a\{t\}+b\{t\}$

where:

$b\{t\}=(S'\{t\}-S''\{t\})$

$a\{t\}= 2S'\{t\}-S''\{t\}$

$S''\{t\}=[S'\{t\}+S'\{t-1\}+S'\{t-2\}]/3$

$S'\{t\}=[X\{t\}+X\{t-1\}+X\{t-2\}]/3$

Initialization:

$S'\{3\}=[X\{3\}+X\{2\}+X\{1\}]/3$

$S'\{4\}=[X\{4\}+X\{3\}+X\{2\}]/3$

$S'\{5\}=[X\{5\}+X\{4\}+X\{3\}]/3$

$S''\{5\}=[S'\{5\}+S'\{4\}+S'\{3\}]/3$

$a\{5\}=2S'\{5\}-S''\{5\}$

$b\{5\}=S'\{5\}-S''\{5\}$

$F\{6\}=a\{5\}+b\{5\}$

Note: This method employs double smoothing. Three sets of three consecutive months are averaged, and the resulting three averages are averaged to create a double-smoothed effect.

3. Four-Month Linear Moving Average-Lag One Month

$F\{t+1\}=a\{t\}+b\{t\}$

$b\{t\}=2/3\,(S'\{t\}-S''\{t\})$

$a\{t\}=2S'\{t\}-S''\{t\}$

$S''\{t\}=[S'\{t\}+S'\{t-1\}+S'\{t-2\}+S'\{t-3\}]/4$

$S'\{t\}=[X\{t\}+X\{t-1\}=X\{t-2\}+X\{t-3\}]/4$

Initialization:

$S'\{4\}=[X\{4\}+X\{3\}+X\{2\}+X\{1\}]/4$

$S'\{5\}=[X\{5\}+X\{4\}+X\{3\}+X\{2\}]/4$

$S'\{6\}=[X\{6\}+X\{5\}+X\{4\}+X\{3\}]/4$

$S'\{7\}=[X\{7\}+X\{6\}+X\{5\}+X\{4\}]/4$

$S''\{7\}=[S'\{7\}+S'\{6\}+S'\{5\}+S'\{4\}]/4$

$a\{7\}=2S'\{7\}-S''\{7\}$

$b\{7\}=2/3\,(S'\{7\}-S''\{7\})$

$F\{8\}=a\{7\}+b\{7\}$

Note: This method smooths the four preceeding monthly observations to create four averages. These four averages are averaged themselves.

4. Two-Month Linear Moving Average-Lag Twelve Months

$F\{t+12\}=a\{t\}+b\{t\}$

$b\{t\}=2(S'\{t\}-S''\{t\})$

$a\{t\}=2S'\{t\}-S''\{t\}$

$S''\{t\}=[S'\{t\}+S'\{t-12\}]/2$

$S'\{t\}=[X\{t\}+X\{t-12\}]/2$

Initialization:

$S'\{13\}=[X\{13\}+X\{1\}]/2$

$S'\{25\}=[X\{25\}+X\{13\}]/2$

$S''\{25\}=[S'\{25\}+S'\{13\}]/2$

$a\{25\}=2S'\{25\}-S''\{25\}$

$b\{25\}=2(S'\{25\}-S''\{25\})$

$F\{37\}=a\{25\}+b\{25\}$

Note: By using observations that are one year apart, seasonality is eliminated. This method is identical to the Two-Month Linear Moving Average above except that the data used in this method are lagged one year.

C. Single Exponential Smoothing

Eighteen variations of the single exponential smoothing technique were tried. The first nine variations are based on shifting the alpha weight from 0.1 to 0.9 by 0.1. The second nine variations incorporated the shifting alpha weight and employ data that are lagged one year.

1. Single Exponential Smoothing (the first nine variations)

F{t+1}=F{t}+alpha (X{t}-F{t})

where alpha steps by 0.1 from 0.1 to 0.9.

Initialization:

F{2}=X{1}

F{3}=X{1}+alpha (X{2}-X{1})

Note: The difference between the last forecast and observation are given different weights in determining the next forecast. The gap between forecast and observation contributes the least when alpha is 0.1 and makes a large contribution when alpha is 0.9.

2. Single Exponential Smoothing (the second nine variations)

F{t+12}=F{t}+alpha (X{t}-F{t})

where alpha steps by 0.1 from 0.1 to 0.9

Initialization:

F{13}=X{1}

F{25}=X{1}+alpha (X{13}-X{1})

Note: The data are lagged one year to minimize the effects of seasonality.

D. Brown's One-Parameter Linear Exponential Smoothing

Eighteen variations of Brown's One-Parameter Linear Exponential Smoothing were tried. The first nine variations are based on shifting the alpha weight from 0.1 to 0.9 by 0.1. The second nine variations incorporated the shifting weight and employed data that are lagged one year.

1. Brown's One-Parameter Linear Exponential Smoothing (the first nine variations)

F{t+1}=a{t}+b{t}m

where m = 1

b{t}=[alpha/(1-alpha)] (S'{t}-S"{t})

where alpha steps by 0.1 from 0.1 to 0.9

a{t}=2S'{t}-S"{t}

S"{t}=alpha S'{t}+(1-alpha)S"{t-1}

S'{t}=alpha X{t}+(1-alpha)S'{t-1}

Initialization:

S'{1}=X{1}

S"{1}=X{1}

S'{2}=alpha X{2}+(1-alpha)X{1}

S"{2}=alpha S'{2}+(1-alpha)X{1}

a{2}=2S'{2}-S"{2}

b{2}=alpha/(1-alpha) (S'{2}-S"{2})

F{3}=a{2}+b{2}

Note: Brown's One-Parameter Linear Exponential Smoothing method is a composite of the Linear Moving Average method and the Single Exponential Smoothing method. As is true with the exponential smoothing methods, heavier weights will make this method respond more quickly to rapidly fluctuating seasonal data.

2. Brown's One-Parameter Linear Exponential Smoothing (the second nine variatons)

F{t+12}=a{t}+b{t}m

where m = 1

b{t}=[alpha/(1-alpha)](S'{t}-S"{t})

where alpha steps by 0.1 from 0.1 to 0.9

a{t}=2S'{t}-S"{t}

S'{t}=alpha S'{t}+(1-alpha)S"{t-12}

S"{t}=alpha X{t}+(1-alpha)S'{t-12}

Initialization:

S'{1}=X{1}

S"{1}=X{1}

S'{13}=alpha X{13}+(1-alpha)X{1}

S"{13}=alpha S'{13}+(1-alpha)X{1}

a{13}=2S'{13}-S"{13}

b{13}=[alpha/(1-alpha)](S'{13}-S"{13})

F{25}=a{13}+b{13}

Note: Every twelfth observation is used to minimize the effect of seasonality.

E. Brown's Quadratic Exponential Smoothing

Eighteen variations of Brown's Quadratic Exponential Smoothing were tested. The first nine variations are based on shifting the alpha weight from 0.1 to 0.9 by 0.1. The second nine variations incorporated the shifting weight and employed data that are lagged one year.

1. Brown's Quadratic Exponential Smoothing

(the first nine variations)

F{t+1}=a{t}+b{t}m + 1/2 c{t} (m squared)

where m = 1

c{t}=[alpha squared/(1-alpha)squared]

(S'{t}-2S"{t}+S'''{t})

where alpha stepped by 0.1 from 0.1 to 0.9

b{t}=[alpha/2(1-alpha)squared]

[(6-5alpha)S'{t}-(10-8alpha)S"{t}+

(4-3alpha)S'''{t}]

a{t}=3S'{t}-3S"{t}+S'''{t}

S'''{t}=alpha S"{t}+(1-alpha)S'''{t-1}

S"{t}=alpha S'{t}+(1-alpha)S"{t-1}

S'{t}=alpha X{t}+(1-alpha)S'{t-1}

Initialization:

S'{1}=X{1}

S"{1}=X{1}

S'''{1}=X{1}

S'{2}=alpha X{2}+(1-alpha)X{1}

S"{2}=alpha S'{2}+(1-alpha)X{1}

S'''{2}=alpha S"{2}+(1-alpha)X{1}

a{2}=3S'{2}-3S"{2}+S"{2}

b{2}=[alpha/2(1-alpha)squared]

[(6-5alpha)S'{2}-(10-8alpha)S"{2}+

(4-3alpha)S'''{2}

c{2}=[alpha squared/(1-alpha)squared]

(S'{2}-2S"{2}+S'''{2})

F{3}=a{2}+b{2}+1/2 c{2}

Note: Brown's quadratic method, which is an extension of linear exponential smoothing, focuses on trends that are more complex than linear trends. Triple exponential smoothing is used. As with previous weighting methods, heavier weights allow this method to track rapidly fluctuating data.

2. Brown's Quadratic Exponential Smoothing

(the second nine variations)

F{t+12}=a{t}m + b{t}+1/2c{t} (m squared)

where m = 1

c{t}=[alpha squared/(1-alpha)squared]

(S'{t}-2S"{t}+S'''{t})

b{t}=[alpha/2(1-alpha)squared]

[(6-5alpha)S'{t}-(10-8alpha)S"{t}+

(4-3alpha)S'''{t}]

a{t}=3S'{t}-3S"{t}+S'''{t}

S'''{t}=alpha S"{t}+(1-alpha)S'''{t-12}

S"{t}=alpha S'{t}+(1-alpha)S"{t-12}

S'{t}=alpha X{t}+(1-alpha)S'{t-12}

Initialization:

S'{1}=X{1}

S"{1}=X{1}

S'''{1}=X{1}

S'{13}=alpha X{13}+(1-alpha)X{1}

S"{13}=alpha S'{13}+(1-alpha)X{1}

S'''{13}=alpha S"{13}+(1-alpha)X{1}

a{13}=3S'{13}-3S"{13}+S'''{13}

b{13}=[alpha/2(1-alpha)squared]

[(6-5alpha)S'{13}-(10-8alpha)S"{13}+

(4-3alpha)S'''{13}]

c{13}=[alpha squared/(1-alpha)squared]

(S'{13}-2S"{13}+S'''{13})

F{25}=a{13}+b{13}+1/2c{13}

Note: Every twelfth observation was used to minimize the effects of seasonality.

F. <u>Holt's Two-Parameter Linear Exponential Smoothing</u>

Eighteen variations of this method were tried. The first nine varied alpha and gamma by 0.4 from 0.1 to 0.9. The two parameters took the values 0.1, 0.5, and 0.9 consecutively resulting in nine combinations. The second nine variations repeated this approach with data lagged one year.

1. Holt's Two-Parameter Linear Exponential Smoothing

(the first nine variations)

$F_{t+1}=S_t+b_t m$

where $m = 1$

$b_t=\text{gamma}\ (S_t-S_{t-1})+\ (1-\text{gamma})\ b_{t-1}$

$S_t=\text{alpha}\ X_t+(1-\text{alpha})\ (S_{t-1}+\ b_{t-1})$

Initialization:

$S_1=X_1$

$b_1=[\ (X_2-X_1)+(X_4-X_3)\ ]/2$

$S_2=\text{alpha}\ X_2+(1-\text{alpha})$

$[\ X_1+((X_2-X_1)+(X_4-X_3))/2\ ]$

$b_2=\text{gamma}\ (S_2-X_1)+$

$(1-\text{gamma})[\ ((X_2-X_1)+(X_4-X_3))/2\ ]$

$F_3=S_2+b_2$

Note: Holt's method is different from preceeding methods because of the necessity of specifying two parameters: alpha and gamma.

2. Holt's Two-Parameter Linear Exponential Smoothing

(the second nine variations)

$F_{t+12}=S_t+b_t m$

where $m = 1$

$b_t=\text{gamma}\ (S_t-S_{t-12})+\ (1-\text{gamma})\ b_{t-12}$

$S_t=\text{alpha}\ X_t+(1-\text{alpha})\ (S_{t-12}\ +\ b_{t-12})$

Initialization:

S{1}=X{1}

D{1}=[(X{13}-X{1})+(X{37}-X{25})]/2

S{13}=alpha X{13}+

(1-alpha)[X{1}+((X{13}-X{1})+(X{37}-X{25}))/2]

D{13}=gamma (S{13}-X{1})+

(1-gamma)[((X{13}-X{1})+(X{37}-X{25}))/2]

F{25}=S{13}+D{13}

Note: The effects of seasonality are controlled by using data lagged one year.

G. Adaptive-Response-Rate Single Exponential Smoothing

Nine variations of this method were tried. Beta was varied from 0.1 to 0.9 by 0.1.

F {t+1}=alpha {t}X{t} + (1-alpha{t})F{t}

Where alpha {t} = |E{t}/M{t}|

E{t} = beta e{t} + (1-beta)E{t-1}

M{t}=beta|e{t}| + (1-beta)M{t-1}

e{t} = X{t} - F {t}

Initialization:

F{2}=X{1}

e{2}=X{2}-X{1}

M{2}=beta|X{2}-X{1}|

E{2}=beta(X{2}-X{1})

alpha {2}=beta

F{3}=beta X{2} + (1-beta)X{1}

e{3} = X{3}-F{3}

M{3}=beta|X{3}-F{3}|

E{3}=beta(X{3}-F{3})

alpha{3}=beta

F{4}=betaX{3} + (1-beta)F{3}

Note: This method requires only the specification of the beta value. This method is adaptive in the sense that the alpha value will change when there is a change in the basic pattern of the data.

H. <u>Winters' Linear and Seasonal Exponential Smoothing</u>

Winters' method works on the development of three equations. Each equation focuses on one aspect of a pattern in time-series data: the stationary level, the trend, and the seasonality of the data. The values of the parameters were symmetrically varied from 0.1 to 0.7 by 0.1. That is, the parameters were assigned the same value so that, for example, alpha = gamma = beta = 0.1. Later, other variations were tried in an effort to achieve lower errors. Table II shows the specification of the parameters used.

$$F\{t+m\}=(S\{t\} + b\{t\}M)\ I\ \{t-L+m\}$$

where L, the lag, is one year

$$F\{t+1\}=(S\{t\} + b\{t\})\ I\{t-12\}$$

$$S\{t\}=(alpha(X\{t\}/I\{t-L\})) + ((1-alpha)(s\{t-1\}+b\{t-1\}))$$

$$b\{t\}=gamma(S\{t\}-S\{t-1\}) + (1-gamma)b\{t-1\}$$

$$I\{t\}=(beta(X\{t\}/S\{t\})) + (1-beta)I\{t-L\}$$

Initialization:

$$S\{13\}=X\{13\}$$

$$b\{13\}=[(x\{13\}-x\{1\}) + (x\{14\}-x\{2\}) + (x\{15\}-x\{3\})]/36$$

$$I\{1\}=X\{1\}/X$$

:

:

$$I\{12\}=X\{12\}/X$$

where $X=sum\ X\{1\ to\ 13\}/13$

$$F\{14\}=(X\{13\} + \{13\})\ I\{2\}$$

I. A Typical Trend Equation

The following method was used as a general purpose equation for trend.

$$F\{t+1\} = 2X\{t\}-X\{t-1\}$$

Note: This is not a smoothing equation described by Makridakis and Wheelwright but is based on our observation of the reducibility of smoothing methods into several generic equations. See Appendix 2 for an exposition of the reducibility of smoothing forecasting methods.

RESULTS

The results of this study are presented in Table I through Table IV.

Analysis of Tables I, II, and III

The first three tables are headed by columns for 1. average percentage error, 2. standard deviation (SD), 3. coefficient of variation (CV), and 4. a minimization of the mean and standard deviation, hereafter called average forys number (AFN).

The average percentage error is the percent error in forecasting after the indicated number of runs averaged for all 50 libraries. In other words, all the libraries are treated as if they were fifty variations of one library and the forecasting results were averaged for all fifty libraries. A measure of variability for this mean is the standard deviation (SD). The coefficient of variation (CV) is the ratio of standard deviation to the mean. The CV provides a way of comparing the characteristics of the different methods. The average forys number (AFN) is an attempt to provide in one number an index of the size of both the mean error and standard deviation. A good forecasting method would minimize both of these. The Pythagorean theorem provides a method for doing this. The Pythagorean theorem states that the square on the hypotenuse of a right triangle equals the sum of the squares on the other sides. The application of this theorem to create the AFN is based on our following observation. If the average percent error were graphed on one axis and the corresponding standard deviation were graphed on the other axis, each forecasting method could be represented by a plotted point. The single best method would be represented by that point closest to the origin of the axes. The distance any point is from the origin of the axes is the length of the hypothenuse on the graph created by the distances from the origin of the mean and standard deviation. The AFN permits the forecasting methods to be compared. The most efficient forecasting method would minimize both its average percentage error and its standard deviation and thus have the smallest AFN.

Table I presents the results of running the smoothing forecasting methods on the monthly data provided by the fifty libraries. The results are reported after the forty-seventh run through the data so as to facilitate comparison among the methods (i.e., Winters' method absorbed 13 data point for initialization).

Several methods are distinguished in this group with low AFN's. They are: One-Month Moving Average, all of the Single Exponential Smoothing Methods, and the Adaptive-Response-Rate Single Exponential Smoothing (beta = 0.99). As a group the average AFN for the forecasting methods is 46.29. It would appear that simpler forecasting methods have the greater success with monthly data. A real assessment, however, must be made against methods that can handle the seasonality of library data such as Winters' method.

Table II presents the results of running Winters' method on the monthly data supplied by the 50 libraries. The average AFN for Winters' method is 13.33 which is much smaller than the average AFN for the smoothing forecasting methods on monthly data. In fact, the difference in mean AFN is statistically significant ($t = 24.17$, $p = 0.0000$). It is clear that Winters' method minimizes AFN with monthly data. Winters' method takes seasonality into account and therefore outperforms the other smoothing methods with monthly data.

Table III presents the results of running the smoothing forecasting methods on the yearly-lagged data for the 50 libraries. The AFN's seem very much smaller. The average AFN for Table III is 10.63. This is significantly smaller than the average AFN for Table I ($t = 27.16$, $p = 0.0000$) and also significantly smaller than the average AFN for Table II ($t = 3.30$, $p = 0.0036$).

It is clear that smoothing forecasting methods perform much better on yearly-lagged data. The reason is that the yearly-lagged data removes the seasonality from the time series. The de-seasonalized data doesn't stay at the same level each year but remains steady enough for the smoothing methods to do a much better tracking job than the smoothing methods do on strongly seasonal data.

The Two-Month Linear Moving Average has the smallest AFN and may be regarded as the best smoothing forecasting method <u>on the average</u>. It must be emphasized that in Tables I to III that the results are averaged across the 50 libraries for each run. No one particular library can expect to have or did have 3.27 AFN. This AFN is an average for the fifty libraries.

It is premature to conclude that the Two-Month Linear Moving Average should be recommended for all libraries. There is a need to evaluate the many fluctuations of the mean percent error and standard deviation values that occurred during the many forecasts made for each method. Before we recommend any forecasting method, we want to investigate the behavior of each method throughout all its forecasts. For an appreciation of the variability of forecasts other calculations are needed that accummulate forecasting errors through the forecasting cycle for each method. This would give a better picture of the ability of the methods to perform with individual libraries.

Analysis of Table IV

Table IV presents a comparison of the forecasting methods with yearly-lagged data. New comparison calculations have been used: 1. the average Mean Squared Error (MSE), 2. the average Mean Absolute Percentage Error (MAPE), 3. the average Mean Percentage Error (MPE), 4. the average Standard Deviation (SD), and 5. the Average Forys Number (AFN).

Makridakis and Wheelwright define these averages in the following way (p.688-89):

Mean Squared Error (MSE)

The mean squared error is a measure of accuracy computed by squaring the individual error for each item in a data set and then finding the average or mean value of the sum of of those squares. The mean squared error gives greater weight to large errors than to small errors because the errors are squared before being summed.

Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error is the mean or average of the sum of all of the percentage errors for a given data set taken without regard to sign. (That is, their absolute values are summed and the average computed.) It is one measure of accuracy commonly used in quantitative methods of forecasting.

Mean Percentage Error (MPE)

The mean percentage error is the average of all of the percentage errors for a given data set. This average allows positive and negative percentage errors to cancel one another. Because of this, it is sometimes used as a measure of bias in the application of a forecasting method.

Table IV is unique in giving information about the variability in forecasting errors for each forecasting method. The three averages defined above were cumulated throughout the forecasting series for each library for each method and then an average was found. The average of these averages across fifty libraries is presented in Table IV. This table consequently presents average errors, standard deviations, and AFN's that a librarian may expect to find if he were to use any of these methods to forecast his own library's circulation data.

An analysis of Table IV leads to the following ranking of smoothing forecasting methods with yearly-lagged data:

1) One-Month Single Moving Average

This ranks first because it had the smallest AFN, the smallest average MSE, the second smallest MAPE, and the second smallest MPE. If we quantify this performance and give a score of one for having the smallest mean, a two for having the second smallest mean, etc., then this method has a comparative rank score of 1 + 1 + 2 + 2 = 6.

2) Brown's One-Parameter Linear Exponential Smoothing (alpha = 0.5)

This method ranks second because it had the second smallest average MSE, the second smallest average AFN, and the third smallest average MAPE and fourth smallest MPE. This method has a comparative rank score of 2 + 2 + 3 + 4 = 11.

3) Single Exponential Smoothing (alpha = 0.99)

This method ranks third because it had the smallest average MAPE, the third smallest average AFN and MSE, and the sixth smallest average MPE. Thus this method has a comparative rank score of 1 + 3 + 3 + 6 = 13.

4) Brown's Quadratic Exponential (alpha = 0.3)

This method ranks fourth because it had the fourth largest MSE, MAPE and AFN and the fifth largest MPE. Thus its comparative rank score is 4 + 4 + 4 + 5 = 17.

5) Holt's Two-Parameter Linear Exponential Smoothing (a = 0.9, g = 0.5)

This method ranks fifth because it scored fifth on average AFN, MSE, and MAPE. It ranked third with MPE. Thus its comparative rank score is 5 + 5 + 5 + 3 = 18.

6) Two-Month Linear Moving Average

This method ranks sixth as it had the seventh largest MSE, MAPE, AFN but the smallest MPE. Thus its comparative rank score is 7 + 7 + 7 + 1 = 22.

7) Adaptive-Response-Rate Single Exponential Smoothing (beta = 0.99)

This method ranks last because it had the sixth highest mean scores for all the averages calculated except MPE where it ranked last. Thus its comparative rank score is 6 + 6 + 6 + 7 = 25.

TABLE I

RESULTS OF USING MONTHLY DATA

	Average % Error after 47 Runs	SD	CV	AFN*
Single Moving Averages				
One-Month	16.04	35.62	2.22	39.06
Two-Month	20.08	38.53	1.92	43.45
Three-Month	22.77	36.16	1.59	42.73
Four-Month	25.13	34.98	1.39	43.07
Linear Moving Averages				
Two-Month	11.65	55.18	4.73	56.40
Three-Month	14.07	52.46	3.73	54.31
Four-Month	18.81	46.65	2.48	50.30
Single Exponential Smoothing				
alpha = 0.1	22.97	31.11	1.35	38.67
alpha = 0.2	23.61	31.83	1.37	39.31
alpha = 0.3	22.67	32.4[illegible]	1.43	39.54
alpha = 0.4	21.72	33.05	1.52	39.55
alpha = 0.5	20.62	33.76	1.64	39.56
alpha = 0.6	19.51	34.45	1.77	39.59
alpha = 0.7	18.42	35.02	1.90	39.57
alpha = 0.8	17.37	35.45	2.04	39.48
alpha = 0.9	16.38	35.71	2.18	39.29
alpha = 0.99	15.52	35.83	2.31	39.05
Brown's One-Parameter Linear Exponential Smoothing				
alpha = 0.1	24.27	33.55	1.34	41.41
alpha = 0.2	22.64	33.96	1.50	40.81
alpha = 0.3	20.20	36.22	1.79	41.47
alpha = 0.4	17.50	39.05	2.23	42.79
alpha = 0.5	15.04	41.65	2.77	44.28
alpha = 0.6	12.87	43.51	3.38	45.37
alpha = 0.7	10.93	44.60	4.08	45.92
alpha = 0.8	9.22	45.26	4.91	46.19
alpha = 0.9	7.65	45.96	6.01	46.59
alpha = 0.99	6.28	47.02	7.48	47.44

*An explanation of these abbreviations used in Tables I through III is given in the Results section.

TABLE I (Con't.)

Brown's Quadratic Exponential Smoothing				
alpha = 0.1	23.61	33.53	1.42	41.01
alpha = 0.2	20.65	36.76	1.78	42.16
alpha = 0.3	16.13	42.09	2.61	45.07
alpha = 0.4	12.27	47.13	3.84	48.70
alpha = 0.5	9.97	49.06	4.92	50.06
alpha = 0.6	6.75	52.25	7.74	52.65
alpha = 0.7	4.58	54.01	11.79	54.20
alpha = 0.8	2.68	57.49	21.44	57.55
alpha = 0.9	0.78	63.89	81.42	63.89
alpha = 0.99	-1.26	72.66	-57.66	72.67
Holt's Two-Parameter Linear Exponential Smoothing				
alpha = 0.1 gamma = 0.1	66.23	66.89	1.01	94.13
alpha = 0.1 gamma = 0.5	32.74	45.18	1.38	55.80
alpha = 0.1 gamma = 0.9	28.21	39.21	1.29	48.30
alpha = 0.5 gamma = 0.1	27.66	37.34	1.35	46.47
alpha = 0.5 gamma = 0.5	18.11	41.65	2.30	45.42
alpha = 0.5 gamma = 0.9	15.41	51.01	3.31	53.29
alpha = 0.9 gamma = 0.1	19.34	38.09	1.97	42.72
alpha = 0.9 gamma = 0.5	11.82	44.20	3.74	45.75
alpha = 0.9 gamma = 0.9	8.27	47.26	5.71	47.98
alpha = 0.99 gamma = 0.99	6.38	47.13	7.39	47.56
Adaptive-Response-Rate Single Exponential Smoothing				
beta = 0.1	27.59	32.56	1.16	42.68
beta = 0.2	25.35	33.71	1.33	42.18
beta = 0.3	22.17	34.80	1.57	41.26
beta = 0.4	21.15	36.37	1.72	42.07
beta = 0.5	20.05	37.89	1.89	42.87
beta = 0.6	18.41	38.67	2.10	42.83
beta = 0.7	17.06	38.38	2.25	42.00
beta = 0.8	16.19	37.88	2.34	41.19
beta = 0.9	15.81	37.16	2.35	40.38
beta = 0.99	15.84	35.95	2.27	39.29
Trend Equation	6.13	47.18	7.69	47.58

TABLE II

WINTER'S METHOD ON MONTHLY DATA

	Average % Errors After 47 Runs	SD	CV	AFN
Coefficients = 0.1*	6.02	10.36	1.72	11.98
0.2	6.22	11.24	1.81	12.85
0.3	5.82	10.42	1.79	11.94
0.4	1.98	14.52	7.35	14.65
0.5	3.61	12.15	3.36	12.67
0.6	3.76	11.84	3.15	12.42
0.7	4.37	12.28	2.81	13.04
0.7 0.5 0.5**	4.66	11.58	2.49	12.48
0.9	4.62	17.31	3.75	17.92

*Each of the three parameters took this value (i.e., alpha = 0.1, beta = 0.1, gamma = 0.1)

**The three parameters took these three values (i.e., alpha = 0.7, beta = 0.5, gamma = 0.5)

TABLE III

RESULTS OF USING YEARLY-LAGGED DATA

	Number of Runs	Average % Error after Last Run	SD	CV	AFN
Single Moving Averages					
One-Month	4	5.47	7.40	1.35	9.20
Two-Month	3	7.88	6.09	0.77	9.96
Three-Month	2	12.75	7.10	0.56	14.59
Two-Month Linear Moving Average					
	2	2.33	2.29	0.98	3.27
Single Exponential Smoothing					
alpha = 0.1	3	11.20	9.35	0.83	14.59
alpha = 0.2	3	10.60	8.84	0.83	13.80
alpha = 0.3	3	10.01	8.35	0.83	13.04
alpha = 0.4	3	9.44	7.88	0.83	12.30
alpha = 0.5	3	8.91	7.46	0.84	11.62
alpha = 0.6	3	8.42	7.10	0.84	11.01
alpha = 0.7	3	7.98	6.83	0.86	10.50
alpha = 0.8	3	7.61	6.64	0.87	10.10
alpha = 0.9	3	7.32	6.57	0.90	9.84
alpha = 0.99	3	7.13	6.61	0.93	9.72
Brown's One-Parameter					
alpha = 0.1	3	10.60	8.84	0.83	13.80
alpha = 0.2	3	9.41	7.84	0.83	12.25
alpha = 0.3	3	8.27	6.93	0.84	10.79
alpha = 0.4	3	7.23	6.16	0.85	9.50
alpha = 0.5	3	6.33	5.62	0.89	8.46
alpha = 0.6	3	5.62	5.36	0.95	7.77
alpha = 0.7	3	5.15	5.42	1.05	7.48
alpha = 0.8	3	4.96	5.86	1.18	7.68
alpha = 0.9	3	5.09	6.69	1.31	8.41

TABLE III (Con't.)

Brown's Quadratic Exponential Smoothing					
alpha = 0.1	3	10.00	8.34	0.83	13.02
alpha = 0.2	3	8.23	6.89	0.84	10.73
alpha = 0.3	3	6.62	5.69	0.86	8.73
alpha = 0.4	3	5.27	4.93	0.94	7.22
alpha = 0.5	3	4.30	4.76	1.11	6.41
alpha = 0.6	3	3.83	5.22	1.36	6.47
alpha = 0.7	3	3.96	6.37	1.61	7.50
alpha = 0.8	3	4.82	8.37	1.74	9.66
alpha = 0.9	3	6.51	11.45	1.76	13.17
Holt's Two-Parmeter Linear Exponential Smoothing					
a=0.1 g=0.1	3	14.50	13.40	0.92	19.74
a=0.1 g=0.5	3	13.96	12.77	0.91	18.92
a=0.1 g=0.9	3	13.43	12.15	0.90	18.11
a=0.5 g=0.1	3	10.39	9.14	0.88	13.84
a=0.5 g=0.5	3	8.25	6.87	0.83	10.74
a=0.5 g=0.9	3	6.36	5.04	0.79	8.11
a=0.9 g=0.1	3	7.90	7.30	0.92	10.76
a=0.9 g=0.5	3	5.65	5.61	0.99	7.96
a=0.9 g=0.9	3	4.64	5.66	1.22	7.32
a=0.99 g=0.99	3	5.45	7.67	1.41	9.41
Trend Equation	3	5.60	7.99	1.43	9.76

TABLE IV

COMPARISON OF YEARLY-LAGGED DATA METHODS

	Number of Runs	Average MSE	Average MAPE	Average MPE	Average SD	Average AFN*
Single Moving Averages						
One-Month	4	1382.61	20.38	5.47	34.79	35.56
Two-Month	3	1819.58	22.40	7.88	40.97	41.99
Three-Month	2	4130.92	27.06	12.75	59.00	60.53
Four-Month	1	8119.51	33.38	22.51	87.25	90.11
Two Month Linear Moving Average						
	2	3974.60	29.16	2.33	57.17	57.25
Single Exponential Smoothing						
alpha=0.1	3	5488.67	34.69	11.20	67.59	68.66
alpha=0.2	3	4435.10	31.02	10.60	61.13	62.22
alpha=0.3	3	3690.47	28.13	10.01	55.35	56.43
alpha=0.4	3	2925.95	25.87	9.44	50.25	51.33
alpha=0.5	3	2416.51	23.99	8.91	45.84	46.93
alpha=0.6	3	2039.48	22.48	8.42	42.18	43.26
alpha=0.7	3	1774.04	21.26	7.98	39.28	40.35
alpha=0.8	3	1601.59	20.35	7.61	37.16	38.22
alpha=0.9	3	1507.04	19.73	7.32	35.85	36.89
alpha=0.99	3	1480.68	19.55	7.13	35.39	36.43
Brown's One-Parameter Linear Exponential Smoothing						
alpha=0.1	3	4422.40	30.87	10.60	61.04	62.13
alpha=0.2	3	2893.64	25.51	9.41	50.00	51.08
alpha=0.3	3	1994.06	22.31	8.27	41.84	42.89
alpha=0.4	3	1559.61	20.86	7.23	37.00	37.96
alpha=0.5	3	1448.90	20.46	6.33	35.53	36.33
alpha=0.6	3	1538.76	21.15	5.62	36.72	37.36
alpha=0.7	3	1734.28	23.64	5.15	39.48	40.09
alpha=0.8	3	1993.57	25.54	4.96	42.98	43.62
alpha=0.9	3	2367.14	27.32	5.09	47.21	48.00

*These abbreviations are explained in the Results section.

TABLE IV (Con't.)

Brown's Quadratic Exponential Smoothing

alpha=0.1	3	3563.30	27.81	10.00	55.14	56.23
alpha=0.2	3	2002.27	22.49	8.23	41.94	42.98
alpha=0.3	3	1536.18	21.27	6.62	36.80	37.62
alpha=0.4	3	1725.92	23.14	5.27	39.16	39.68
alpha=0.5	3	2210.21	26.60	4.30	45.00	46.59
alpha=0.6	3	2747.07	29.55	3.83	51.20	51.57
alpha=0.7	3	3346.61	32.49	3.96	56.88	57.41
alpha=0.8	3	4495.16	35.95	4.82	65.06	65.98
alpha=0.9	3	7470.88	44.74	6.51	84.95	86.08

Holt's Two-Parmeter Linear Exponential Smoothing

a=0.1 g=0.1	3	11635.48	39.79	14.50	91.44	92.72
a=0.1 g=0.5	3	10497.98	38.86	13.96	87.80	89.05
a=0.1 g=0.9	3	9424.95	38.07	13.43	84.24	85.46
a=0.5 g=0.1	3	4462.94	31.34	10.39	61.40	62.48
a=0.5 g=0.5	3	2735.62	28.57	8.25	50.12	51.05
a=0.5 g=0.9	3	2240.64	26.81	6.36	45.75	46.38
a=0.9 g=0.1	3	2338.46	26.71	7.90	44.72	45.71
a=0.9 g=0.5	3	1994.47	[illegible].48	5.65	41.72	42.40
a=0.9 g=0.9	3	2434.86	28.02	4.64	47.69	48.23
a=0.99 g=0.99	3	2920.97	29.88	5.45	52.55	53.50

Adaptive-Response-Rate Single Exponential Smoothing

beta = 0.1	1	4894.96	30.90	20.88	66.78	69.96
beta = 0.2	1	4456.54	[illegible]	20.21	63.64	66.77
beta = 0.3	1	4088.80	29.61	19.51	60.89	63.94
beta = 0.4	1	3599.26	28.54	18.60	57.04	59.99
beta = 0.5	1	3176.65	27.41	17.86	53.48	56.38
beta = 0.6	1	2619.02	[illegible]	17.38	50.17	53.09
beta = 0.7	1	2547.08	[illegible].42	16.84	47.58	50.47
beta = 0.8	1	2376.[illegible]	[illegible].09	16.39	45.91	48.75
beta = 0.9	1	2297.94	26.19	16.34	45.07	47.94
beta = 0.99	1	2262.98	26.20	16.32	44.68	47.57

Trend Equation

	3	3051.81	30.42	5.60	53.79	54.79

CONCLUSION

This study has attempted to assess smoothing forecasting methods as efficient predictors of academic library statistics. After examining the results of this study, we would like to make the following points:

1) We don't recommend the application of smoothing forecasting methods on monthly library circulation totals. The data, as exhibited in the plots of Appendix 3, are simply far too seasonal for smoothing methods to model them accurately. This conclusion is in accord with other assessments of these methods, i.e., Makridakis and Wheelwright (1978, p.69) who comment on inability of smoothing methods to handle seasonal data. As revealed in Table II, a method like Winters' which takes seasonality into account does a better forecasting job. But we recommend none of these methods for monthly data. However, if a librarian were to employ smoothing forecasting methods then we would urge him to use Winters' Linear and Seasonal Exponential Smoothing. This method, however, is relatively complicated and consumes alot of data in initialization.

2) We urge the use of smoothing forecasting methods on yearly-lagged data. As had been shown, Table III's AFN scores are significantly smaller than either Tables I and II AFN's.

But before deciding which smoothing forecasting method to use with yearly-lagged data, we investigated the variability of errors that a librarian might expect to find with his library's data. Table IV revealed that the One-Month Moving Average and Brown's One-Parameter Linear Exponential Smoothing (alpha = 0.5) were the two best methods to use on yearly-lagged library circulation totals.

3) We urge academic librarians to regard a plot of monthly circulation data lagged one year and make the following decision: if their data don't show trend across several years, then use the One-Month Moving Average method for predicting future totals. If, on the other hand, their data do trend up or down across several years, then we urge the use of Brown's One-Parameter Linear Exponential Smoothing (alpha = 0.5).

The implications of these recommendations are that with trendless or nearly trendless yearly-lagged data, the best predictor of a month next year is the that month's total this year. With trending yearly-lagged data, the best predictor is a method that smooths several previous yearly-lagged observations together as a trend adjustment.

Our feeling is that more research needs to be done on the pre-analysis stage of forecasting. By examining their data before forecasting, academic librarians may be able to select the most appropriate forecasting method. We believe that improved forecasts will be the result.

BIBLIOGRAPHY

Brooks, Terrence A. 1981. An Analysis Of Library-Output Statistics. Austin, TX: The University of Texas at Austin; 1981. Dissertation.

Burns, Robert W. 1974. An Empirical Rationale For The Accumulation Of Statistical Information. Library Resources Technical Services. 1974 Summer; 18 (3): 253-258.

Burns, Robert W. 1977. Library Performance Measures As Seen In The Statistics Gathered By Automated Circulation Systems. American Society For Information Science, Special Interest Group on Library Automation and Networks, SIG Newsletter. 1977 March; LAN-6: 1-7.

Burrell, Quentin. 1980. A Simple Stochastic Model For Library Loans. Journal of Documentation. 1980 June; 36 (2): 115-132.

Carpenter, Ray L.; Vasu, Ellen S. 1978. Statistical Methods For Librarians. Chicago, IL: American Library Association; 1978.

Chambers, J.C.; Mullick, S.K.; Smith, D.D. 1971. How To Choose The Right Forecasting Technique. Harvard Business Review. 1971 July-August; 49 (4): 45-74.

Dervin, Brenda. 1977. Useful Theory For Librarianship: Communication, Not Information. Drexel Library Quarterly. 1977 July; 13 (3): 16-32.

Drake, Miriam A. 1976. Forecasting Academic Library Growth. College Research Libraries. 1976 January; 37 (1): 53-59.

Hamburg, Morris. 1978. Statistical Methods For Library Management. In Chen, C-C, ed. Quantitative Measurement And Dynamic Library Service. Phoenix, AZ: Oryz Press; 1978.

Herner, Saul. 1967. Meaningful Statistics. In Practical Problems Of Library Automation [papers presented at the 1966- 1967 meetings]. Washington, DC: Documentation Group, Washington Chapter, Special Libraries Association, 1967.

Hoadley, Irene B.; Clark, Alice S., eds. 1972. Quantitative Methods in Librarianship. Westport, CT: Greenwood Press; 1972.

Hodowanec, George V. 1980. Analysis Of Variables Which Help To Predict Book And Periodical Use. Library Aquisitions: Practice and Theory. 1980; 4 (1): 75-85.

Jones, William G. 1973. A Time-Series Sample Approach For Measuring Use In A Small Library. Special Libraries. 1973 July; 64 (7): 280-284.

Kang, Jong Hoa. 1976. Approaches To Forecasting Demands For Library Network Services. Urbana, IL: University of Illinois at Urbana-Champaign; 1979. Dissertation.

Lancaster, F. Wilfrid. 1977. The Measurement And Evaluation Of Library Services. Washington, DC: Information Resources Press; 1977.

Lazorick, Gerald Joseph. 1970. Analysis Of The Stochastic Properties And Prediction Of Demand For Books In Library Circulation Systems. Buffalo, NY: State University of New York at Buffalo; 1970. Dissertation.

McGrath, William E. 1976-77. Predicting Book Circulation By Subject In A University Library. Collection Management. 1976-77 Fall-Winter; 1 (3-4): 7-26.

Makridakis, Spyros; Wheelwright, Steven C. 1978. Forecasting Methods And Applications. New York, NY: John Wiley; 1978.

Martyn, John; Lancaster, F. Wilfrid. 1981. Investigative Methods In Library And Information Science: An Introduction. Arlington, VA: Information Resources Press; 1981.

Morse, Philip M.; Chen, Ching-chih. 1975. Using Circulation Desk Data To Obtain Unbiased Estimates Of Book Use. Library Quarterly. 1975 April; 45 (2): 179-194.

Nozik, Barbara Sayles. 1974. A Stochastic Model To Predict Demand For Library Services. Berkeley, CA: University of California; 1974. Dissertation.

Rogers, Rutherford D.; Weber, David C. 1971. University

Library Administration. New York, NY: H.W. Wilson; 1971.

Rouse, W.B. 1974. Circulation Dynamics: A Planning Model. Journal of the American Society for Information Science. 1974 November-December; 25 (6): 358-363.

Simpson, Ian S. 1975. Basic Statistics For Librarians. London: Clive Bingley; 1975.

Slote, Stanley James. 1970. The Predictive Value Of Past-Use Patterns Of Adult Fiction In Public Libraries For Identifying Core Collections. New Brunswick, NJ: Rutgers University, The State University of New Jersey; 1970. Dissertation.

Stueart, Robert D.; Eastlick, John T. 1981. Library Management. Littleton, CO: Libraries Unlimited; 1981.

Zweizig, Douglas. 1973. Predicting Amount Of Library Use: An Empirical Study Of The Role Of The Public Library In The Life Of The Adult Public. Syracuse, NY: Syracuse University; 1973. Dissertation.

BIOGRAPHIES

Terrence A. Brooks

Terrence Brooks is an Assistant Professor at the School of Library and Information Science, The University of Iowa. He has received the following degrees: B.A. (University of British Columbia, 1968), M.L.S. (McGill University, 1971), M.B.A. (York University, 1975), and Ph.D. (University of Texas at Austin, 1981). He has worked as a librarian at the Halifax City Regional Library, Halifax, Nova Scotia; the El Paso Public Library, El Paso, Texas; and at the University of Iowa, Iowa City, Iowa. His doctoral work concerned the statistical nature of library-output statistics such as monthly circulation data. He teaches the research course at the School of Library Science. He is familiar with four programming languages: BASIC, Fortran, Pascal, and Cobol. He is a member of the American Library Association, the Association of American Library Schools, the Iowa Library Association, the American Society for Information Science, and the International Institute of Forecasters.

John W. Forys, Jr.

John Forys is the Engineering Librarian at the University of Iowa. He has received the following degrees: B.S. in Aerospace Engineering (West Virginia University, 1971) and M.S. in Library Science (University of North Carolina at Chapel Hill, 1976). He has worked as the Assistant Director of the Mary H. Weir Public Library, Weirton, West Virginia; and at the University of Iowa as both the Engineering Librarian and the Engineering/ Mathematics Librarian. He is familiar with three programming languages: BASIC, Fortran, and PL/1 and owns a microcomputer. He is a member of the American Library Association and the Association of College and Research Libraries.

APPENDIX 1

The Letter of Inquiry

In September we began a research project whose aim is to find the most effective method of predicting future circulation levels of academic libraries based on past circulation data. We are writing to you as part of our collection phase; we are looking for academic libraries that would have approximately five years' worth of monthly circulation counts available for analysis in our study.

Our intent is to collect data from about 50 academic libraries in the Midwest. Each library would contribute five years' worth of monthly total circulation counts (i.e., 60 consecutive monthly total circulation counts). Each time series thus collected from each library would then be analyzed with an interactive forecasting software package called SIBYL/RUNNER that is available for research use on a Hewlett Packard 2000 computer here at the University of Iowa. From the output produced by SIBYL/RUNNER we will be able to determine which of 24 extrapolative time-series methods would be able to model each library's data and make the most effective forecasts. Our study of academic library circulation statistics follows in the tradition of other more general forecasting studies, e.g. "The

Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition" by Spyros Makridakis, et al. Journal of Forecasting, v.1, pages 111-153, 1982. This study has pioneered extrapolative techniques in general; we believe we will be the first to apply them to academic library circulation counts.

Can you send us about five years' worth of your total monthly circulation counts? Any consecutive sequence of 60 months in the recent past will do. We know that libraries collect differing statistics and call them different names. These differences will not affect our study for we don't intend to compare the counts from library to library. Instead we want to study the performance of the 24 forecasting methods on many different sets of circulation data.

An example of the type of data we are looking for would be the annual statistical summary many libraries compile giving the total circulation of the main and any branch libraries for each month of the preceeding year. A photocopy of such an annual statistical summary would suit our purposes very well.

The contribution of your library's total monthly counts are important to our study and we would like to thank you now for every effort you expend on our behalf.

APPENDIX 2

The Reducibility of Smoothing Forecasting Methods

Smoothing forecasting methods form a family of methods that, when coefficients are set to 1, reduce down to two simple formulas:

<u>The First Formula:</u>

$F\{t+m\} = X\{t\}$

and when $m = 1$

$F\{t+1\} = X\{t\}$

This is equivalent to saying that a future observation in a time series will be like its immediate predecessor.

<u>The Second Formula:</u>

$F\{t+m\} = X\{t\} + (X\{t\} - X\{t-1\})m$

and when $m = 1$

$F\{t+1\} = X\{t\} + (X\{t\} - X\{t-1\})$

or

$F\{t+1\} = 2X\{t\} - X\{t-1\}$

This is equivalent to saying that a future observation in a time series will be like its immediate predecessor plus the difference observed in the time series between the last two observations.

As is illustrated below, smoothing forecasting methods degenerate into either of these two formulas when coefficients are set to 1.

<u>Single Exponential Smoothing</u>

F{t+1} = alpha X{t} + (1-alpha) F{t}

when alpha = 1

F{t+1} = X{t}

<u>Brown's One-Parameter Linear Exponential Smoothing</u>

S'{t} = alpha X{t} + (1-alpha) S'{t-1}

when alpha = 1

S'{t} = X{t}

S"{t} = alpha S'{t} + (1-alpha) S"{t-1}

when alpha = 1

S"{t} = S'{t}

Therefore S"{t} = S'{t} = X{t}

a{t} = 2X{t} - X{t}

a{t} = X{t}

b{t} = 1/0 (X{t} - X{t})

b{t} = 0

F{t+m} = a{t}

when m = 1 and since a{t} = X{t}

F{t+1} = X{t}

<u>Brown's Quadratic Exponential Smoothing</u>

$S'\{t\} = \text{alpha}X\{t\} + (1-\text{alpha})S'\{t-1\}$

$S''\{t\} = \text{alpha}\ S'\{t\} + (1-\text{alpha})S''\{t-1\}$

$S'''\{t\} = \text{alpha}\ S''\{t\} + (1-\text{alpha})S''\{t-1\}$

when alpha = 1

$S'\{t\} = S''\{t\} = S'''\{t\} = X\{t\}$

$a\{t\} = 3S'\{t\} - 3S''\{t\} + S'''\{t\}$

which reduces to

$a\{t\} = X\{t\}$

$b\{t\} = [\text{alpha}/2\ (1-\text{alpha})\text{squared}]$

$[(6-5\ \text{alpha})\ S'\{t\} - (10-8\ \text{alpha})\ S''t +$

$(4-3\ \text{alpha})\ S'''\{t\}]$

when alpha = 1

$b\{t\} = 0$

$c\{t\} = [\text{alpha squared}/(1-\text{alpha})\text{squared}]$

$(S'\{t\} - 2S''\{t\} + S'''\{t\})$

when alpha = 1

$C\{t\} = 0$

$F\{t+m\} = X\{t\}$

when m = 1

$F\{t+1\} = X\{t\}$

Holt's Two-Parameter Linear Exponential Smoothing

S{t} = alpha X{t} + (1-alpha)(S{t-1} + b{t-1})

when alpha = 1

S{t} = X{t}

b{t} = gamma (S{t} -S{t-1}) + (1-gamma) b{t-1}

when gamma = 1

b{t} = S{t} - S{t-1}

or

b{t} = X{t} - X{t-1}

F{t+m} = S{t} + b{t}m

when m=1

F{t+1} = 2X{t} - X{t-1}

Adaptive-Response-Rate Single Exponential Smoothing

F{t+1} = alpha {t} X{t} + (1-alpha{t}) F{t}

alpha {t+1} = |E{t}/M{t}|

E{t} = beta e{t} + (1-beta)E{t-1}

M{t} = beta |e{t}| + (1-beta)M{t-1}

e{t} = X{t} - F{t}

when beta = 1

E{t} = e{t}

M{t} = |e{t}|

therefore

alpha {t+1} = 1

$F\{t+1\} = X\{t\}$

Winters' Linear and Seasonal Exponential Smoothing

$S\{t\} = \text{alpha } (X\{t\}/I\{t-L\}) + (1-\text{alpha}) (S\{t-1\} + b\{t-1\})$

when alpha = 1

$S\{t\} = X\{t\}/I\{t-L\}$

$b\{t\} = \text{gamma } (S\{t\} - S\{t-1\}) + (1-\text{gamma}) b\{t-1\})$

when gamma = 1

$b\{t\} = S\{t\} - S\{t-1\}$

$I\{t\} = \text{beta } (X\{t\}/S\{t\}) + (1-\text{beta}) I\{t-L\}$

when beta = 1

$I\{t\} = X\{t\}/S\{t\}$

initialize $S\{t\} = X\{t\}$

$I\{t\} = 1$

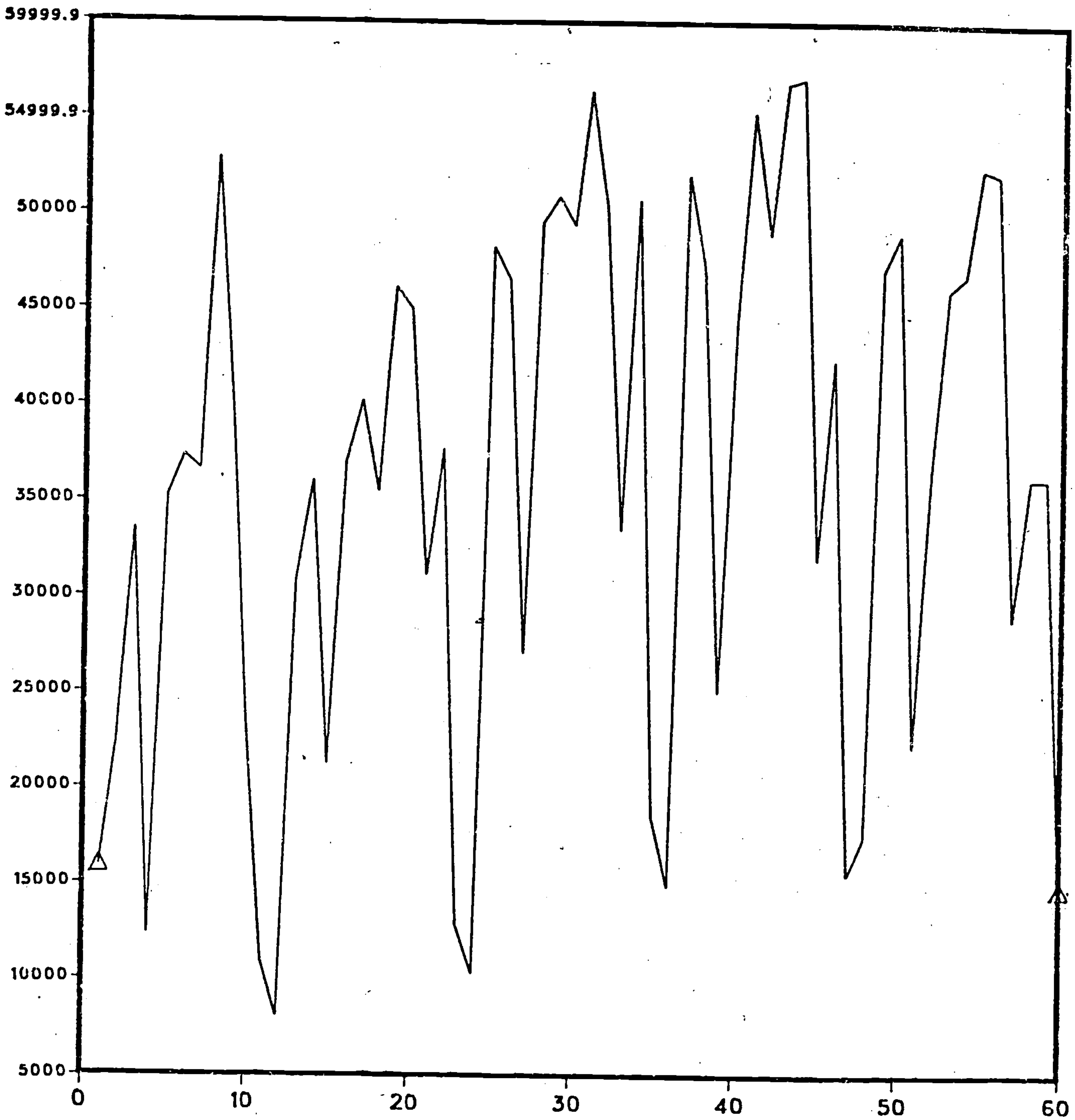
$b\{t\} = X\{t\} - X\{t+1\}$

$F\{t+1\}=2X\{t\} - X\{t-1\}$

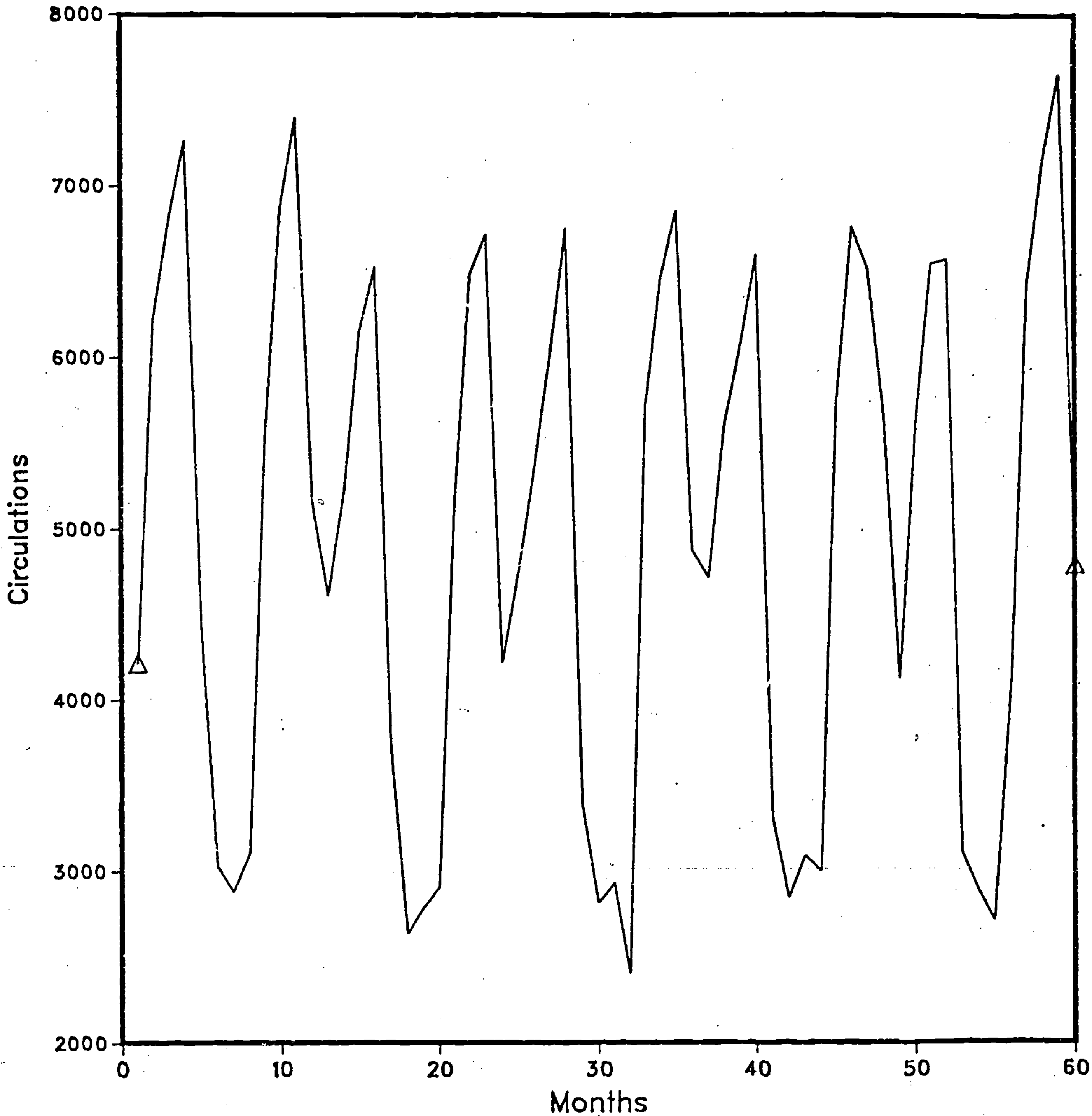
Appendix 3

Circulation Plots of Fifty Libraries

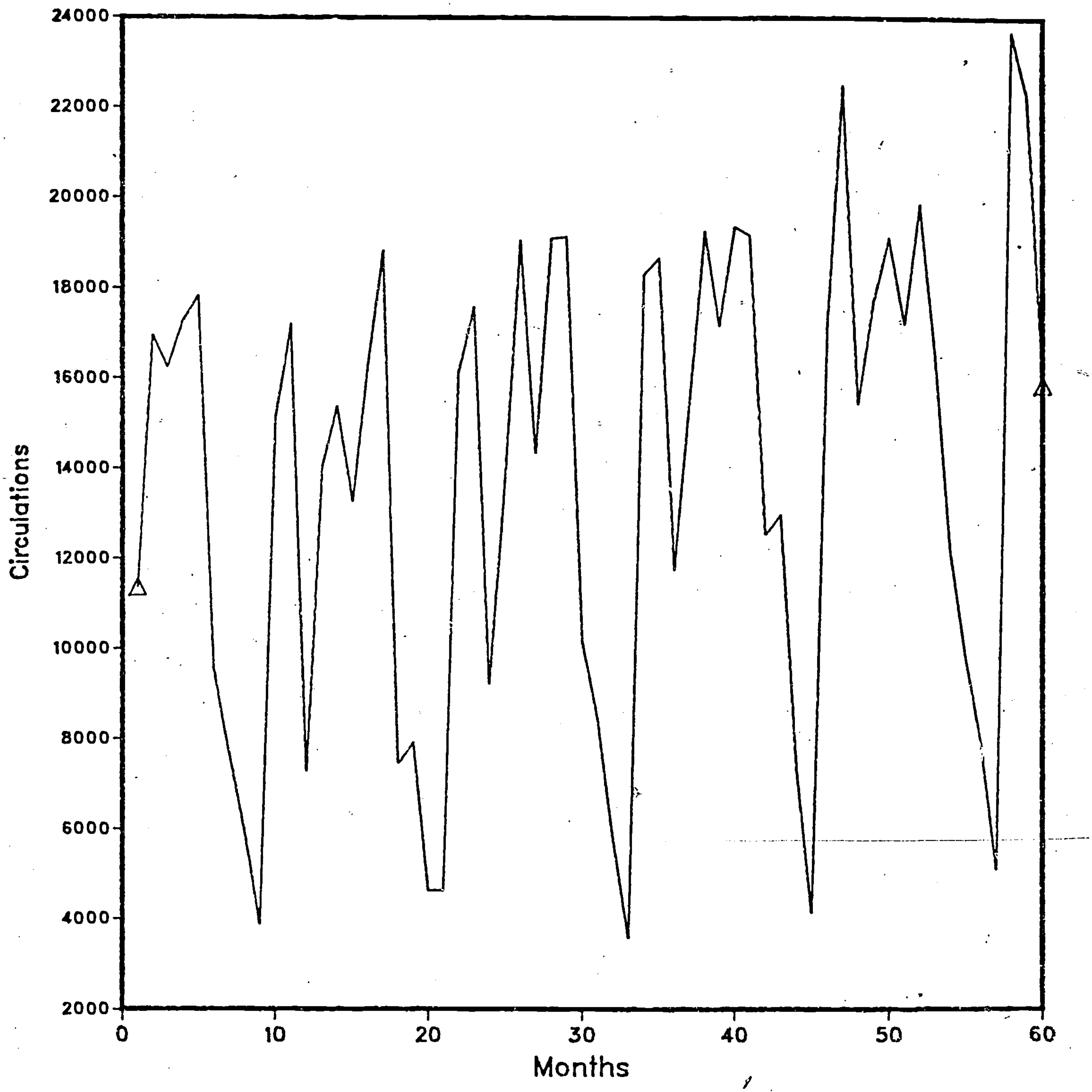
Library 1
Monthly Circulations



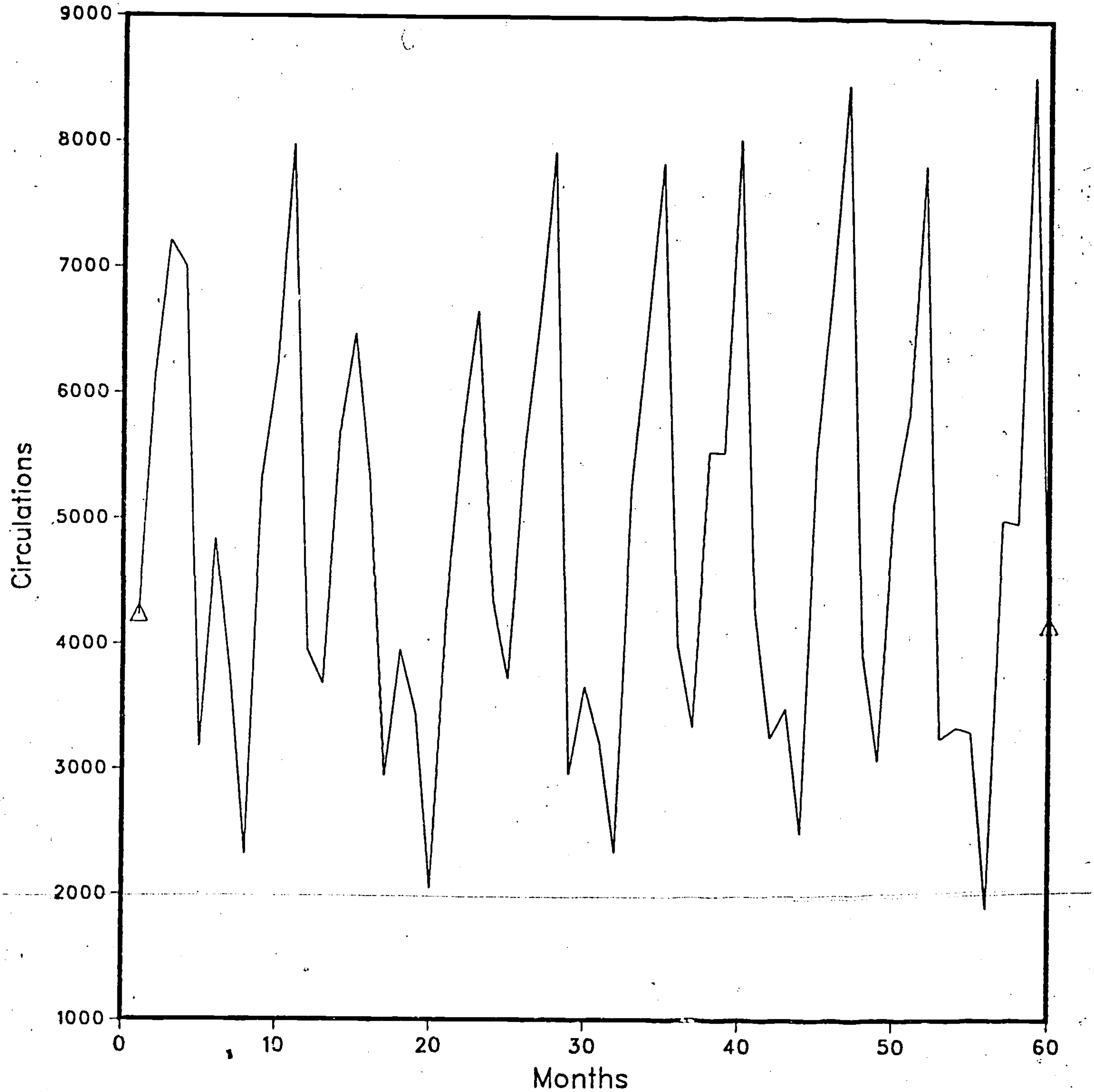
Library 2
Monthly Circulations



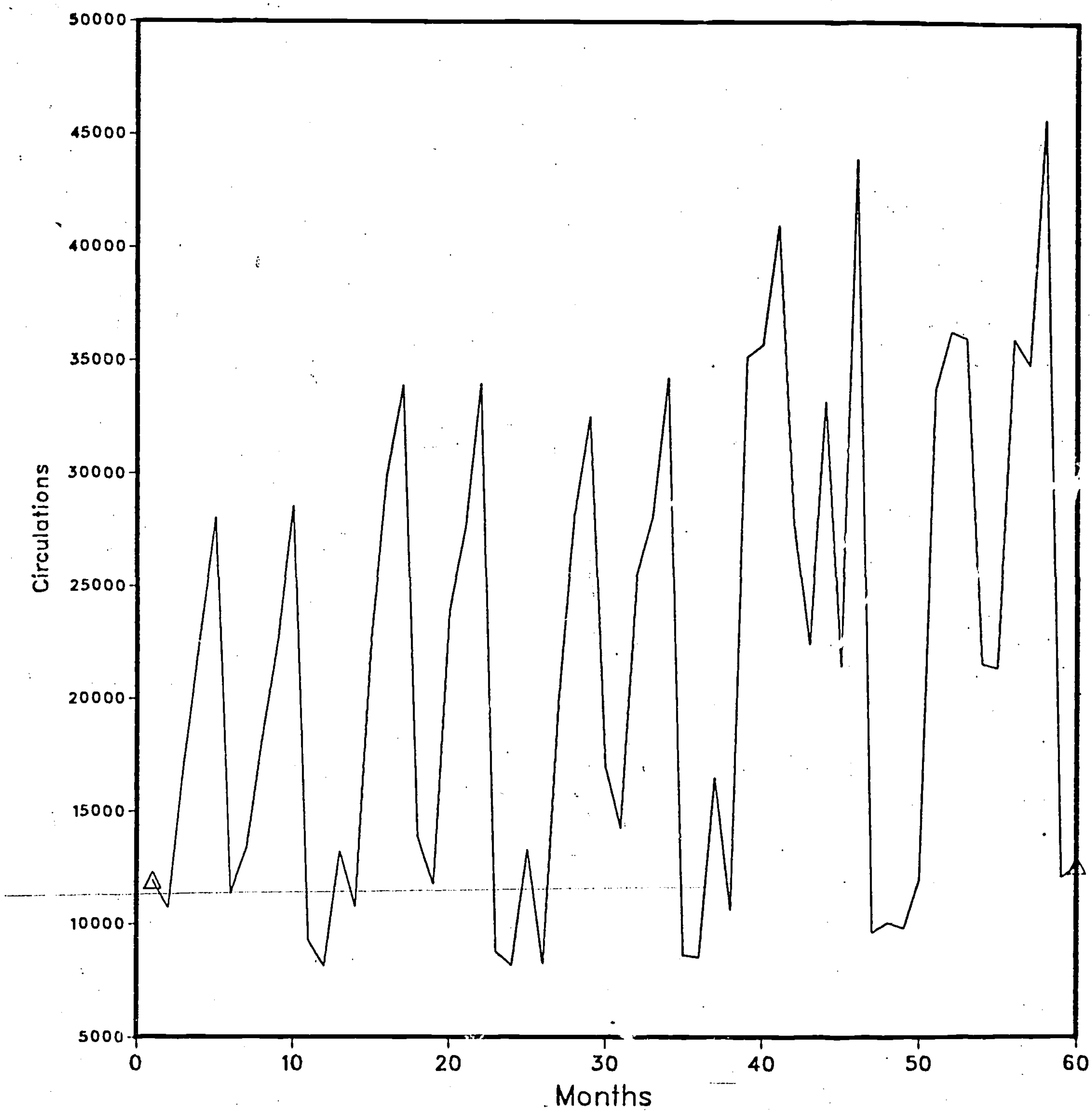
Library 3
Monthly Circulations



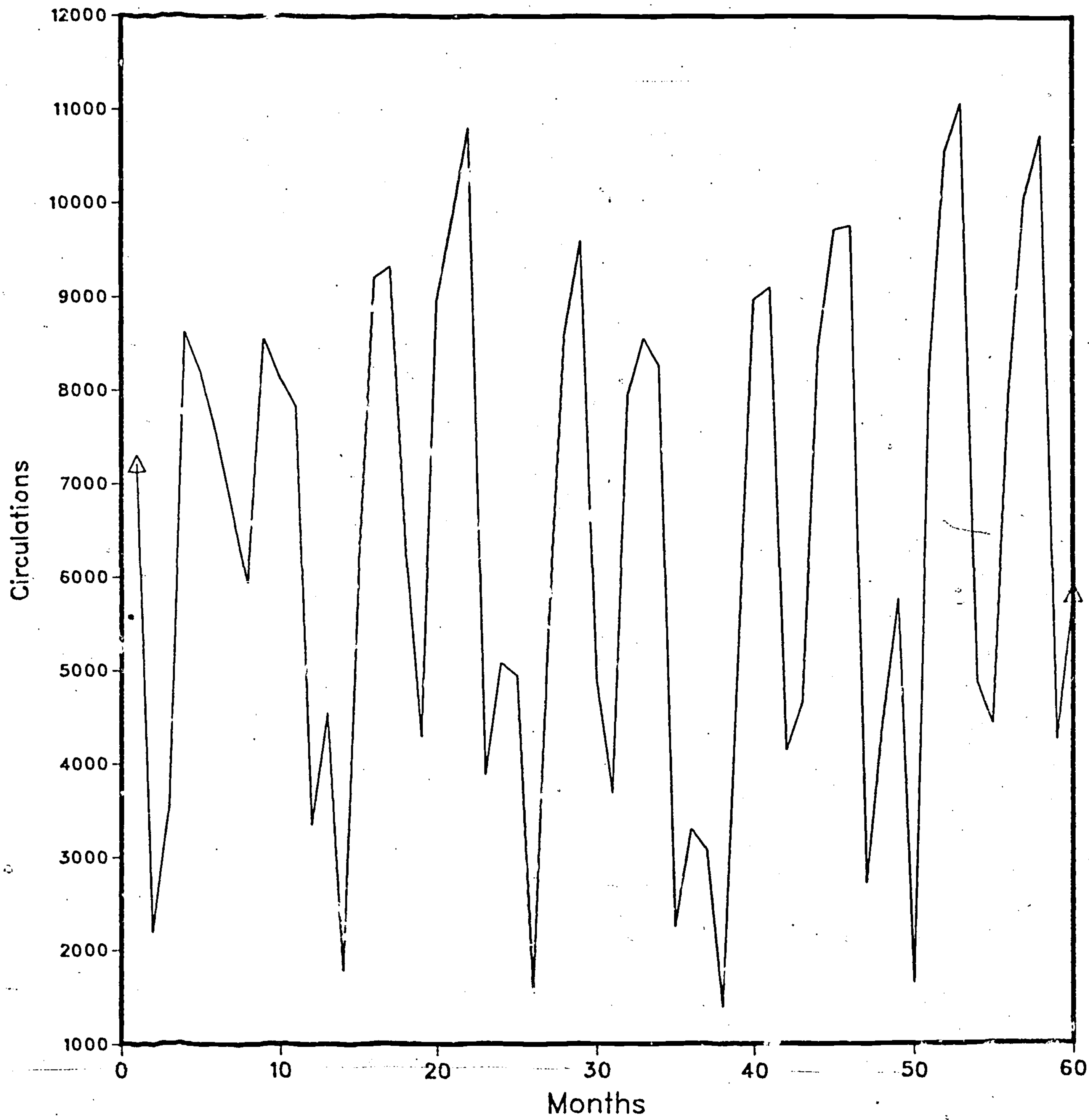
Library 4
Monthly Circulations



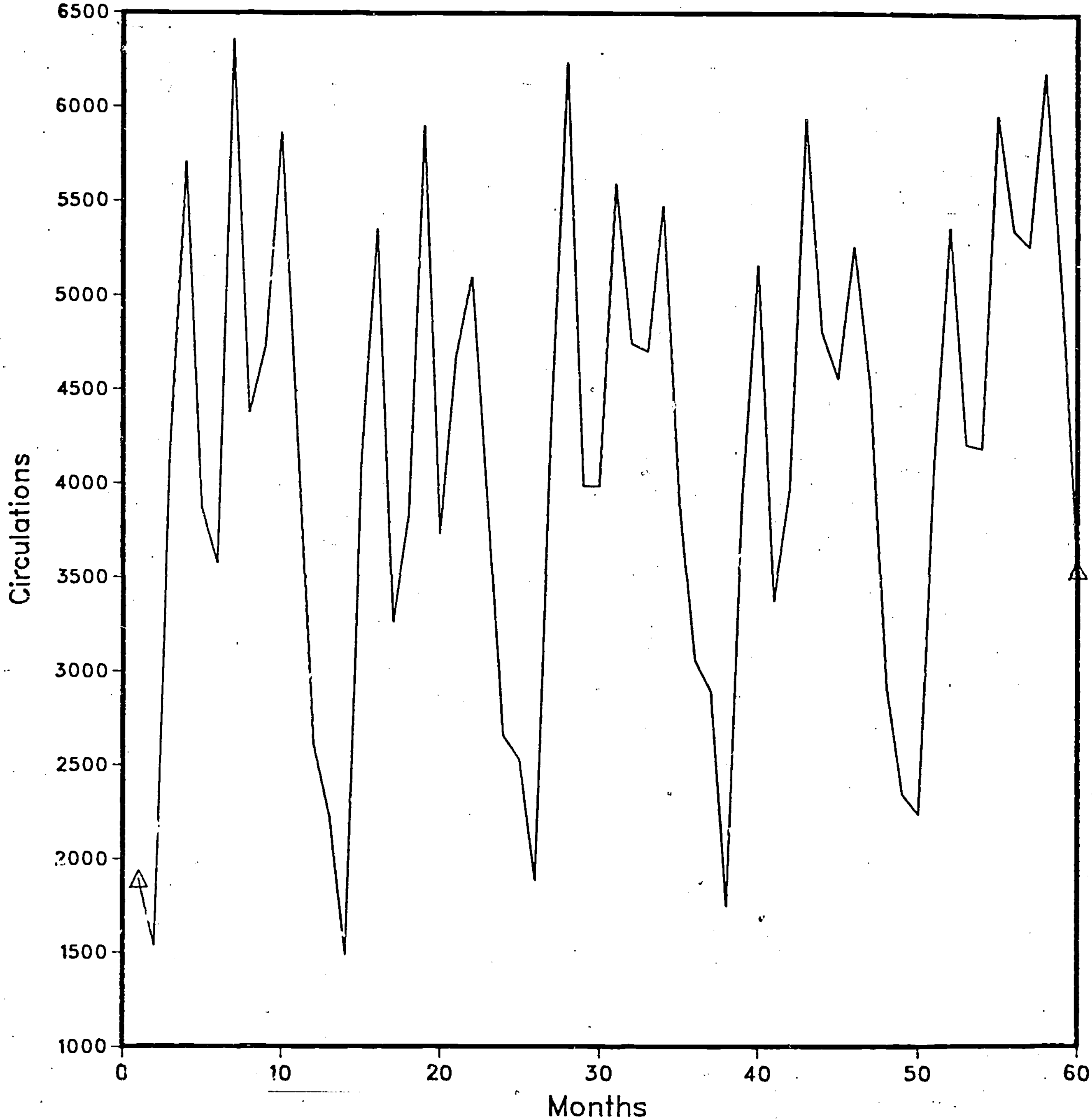
Library 5
Monthly Circulations

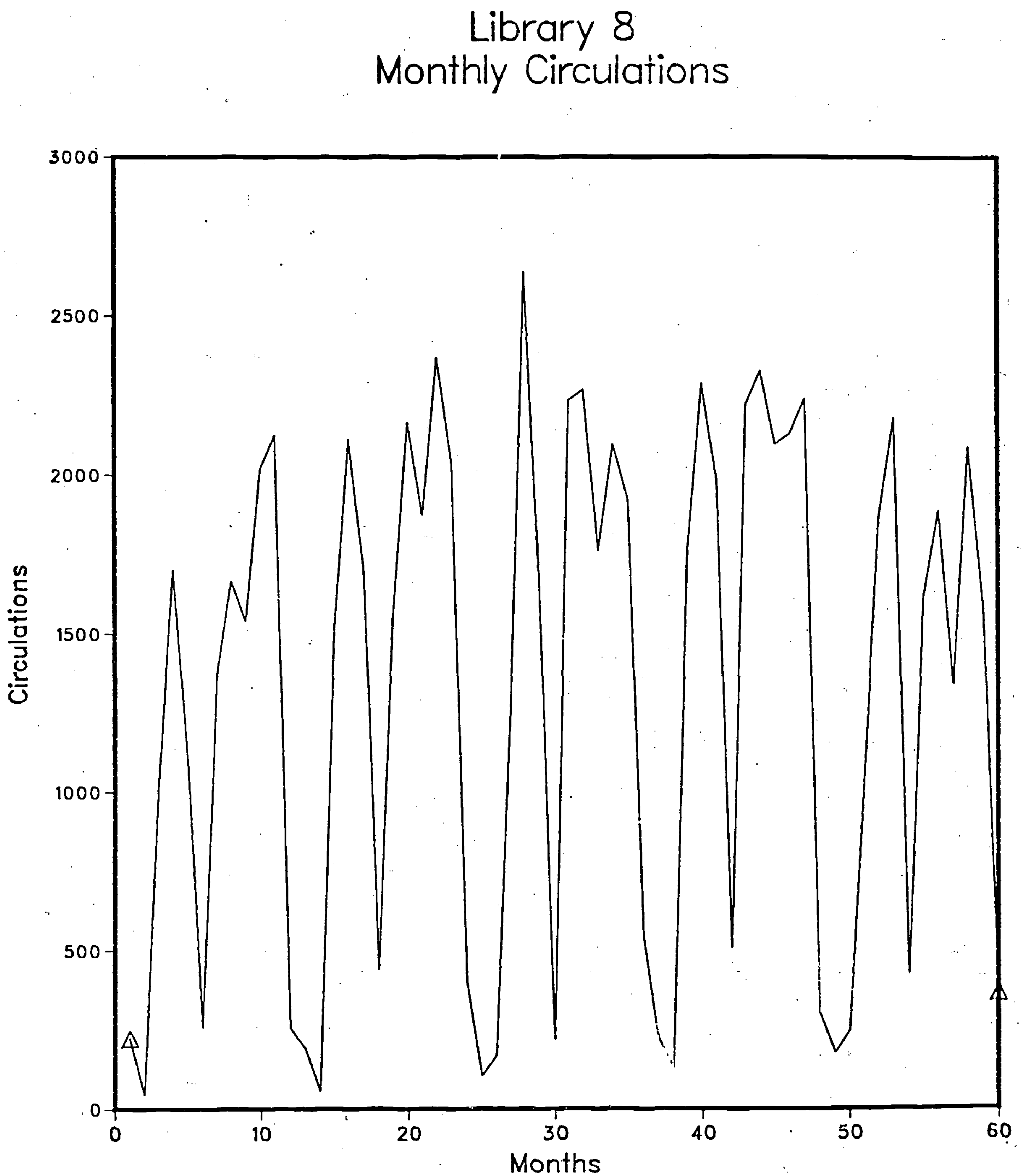


Library 6
Monthly Circulations

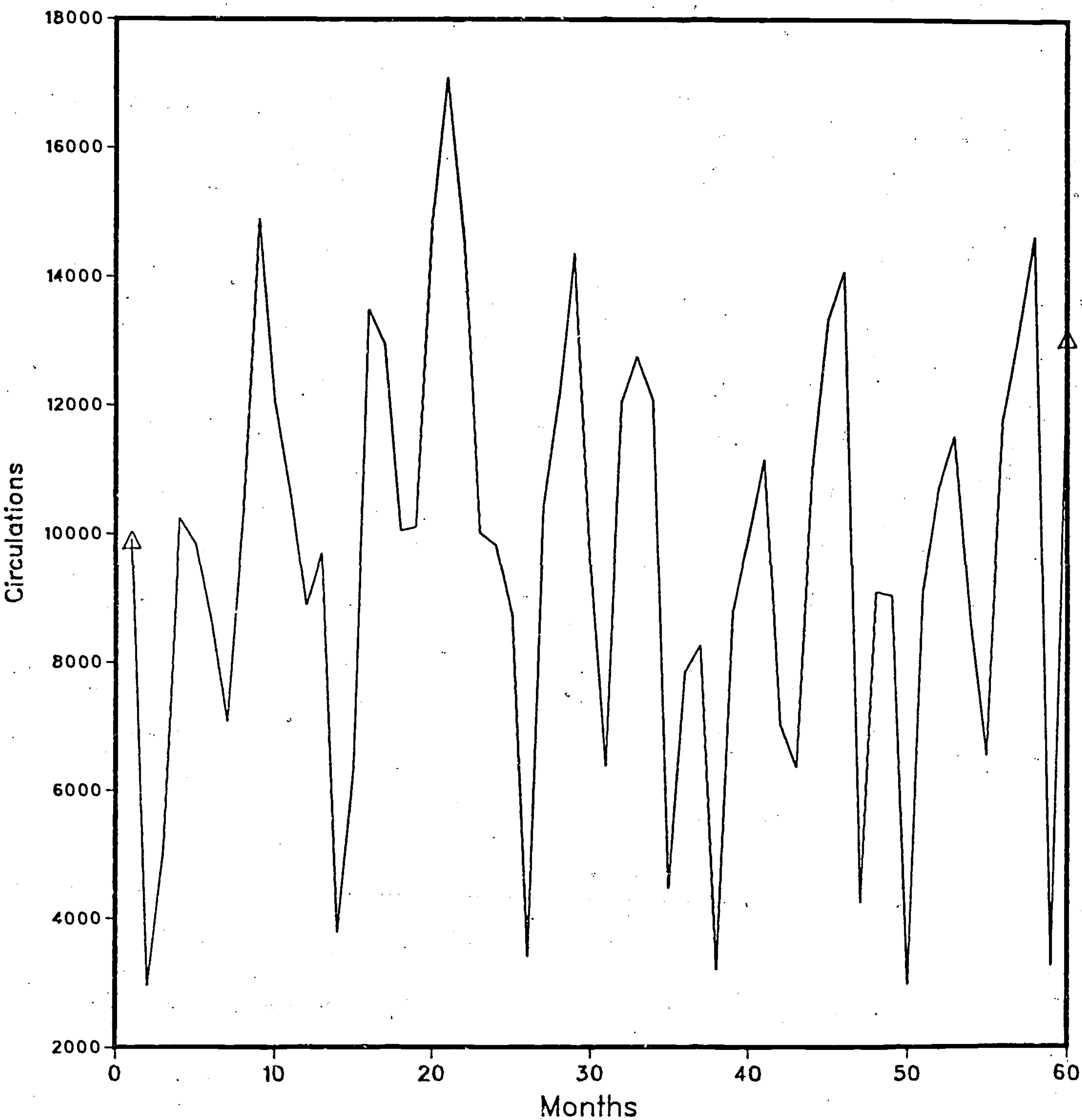


Library 7
Monthly Circulations

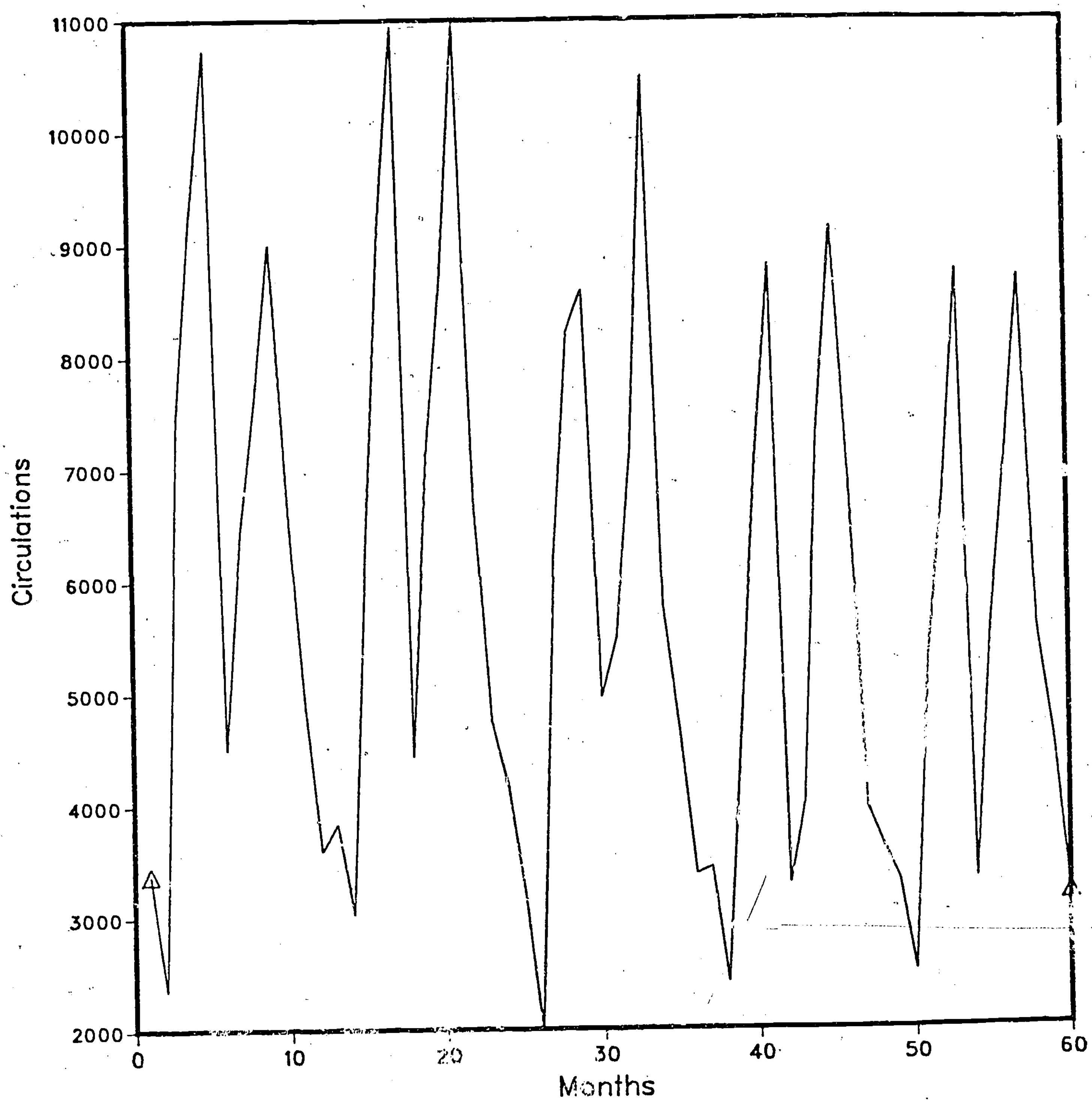



Library 8
Monthly Circulations
Circulations
3000
2500
2000
1500
1000
500
0
0
10
20
30
40
50
60
Months

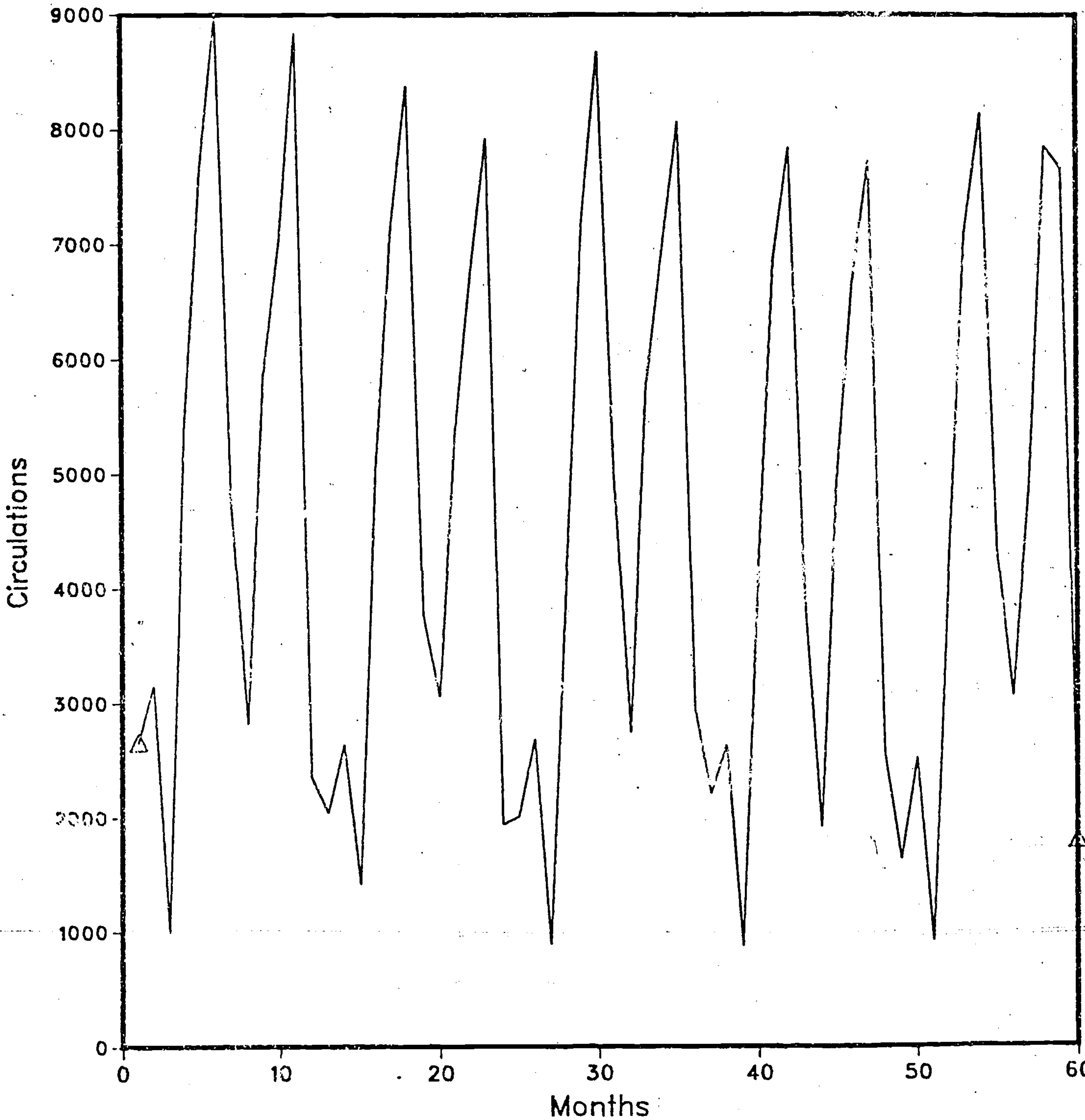
Library 9
Monthly Circulations



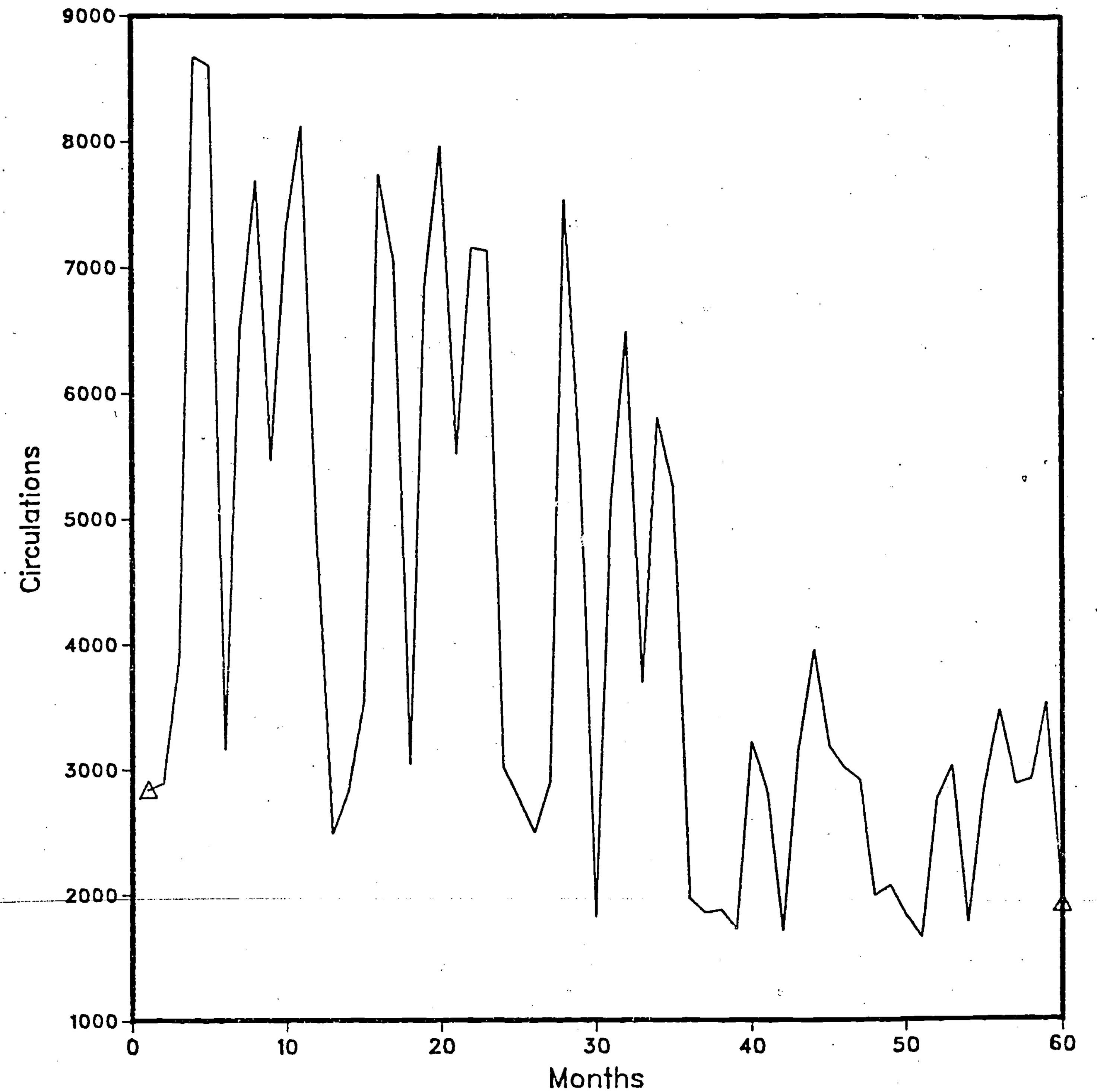
Library 10
Monthly Circulations



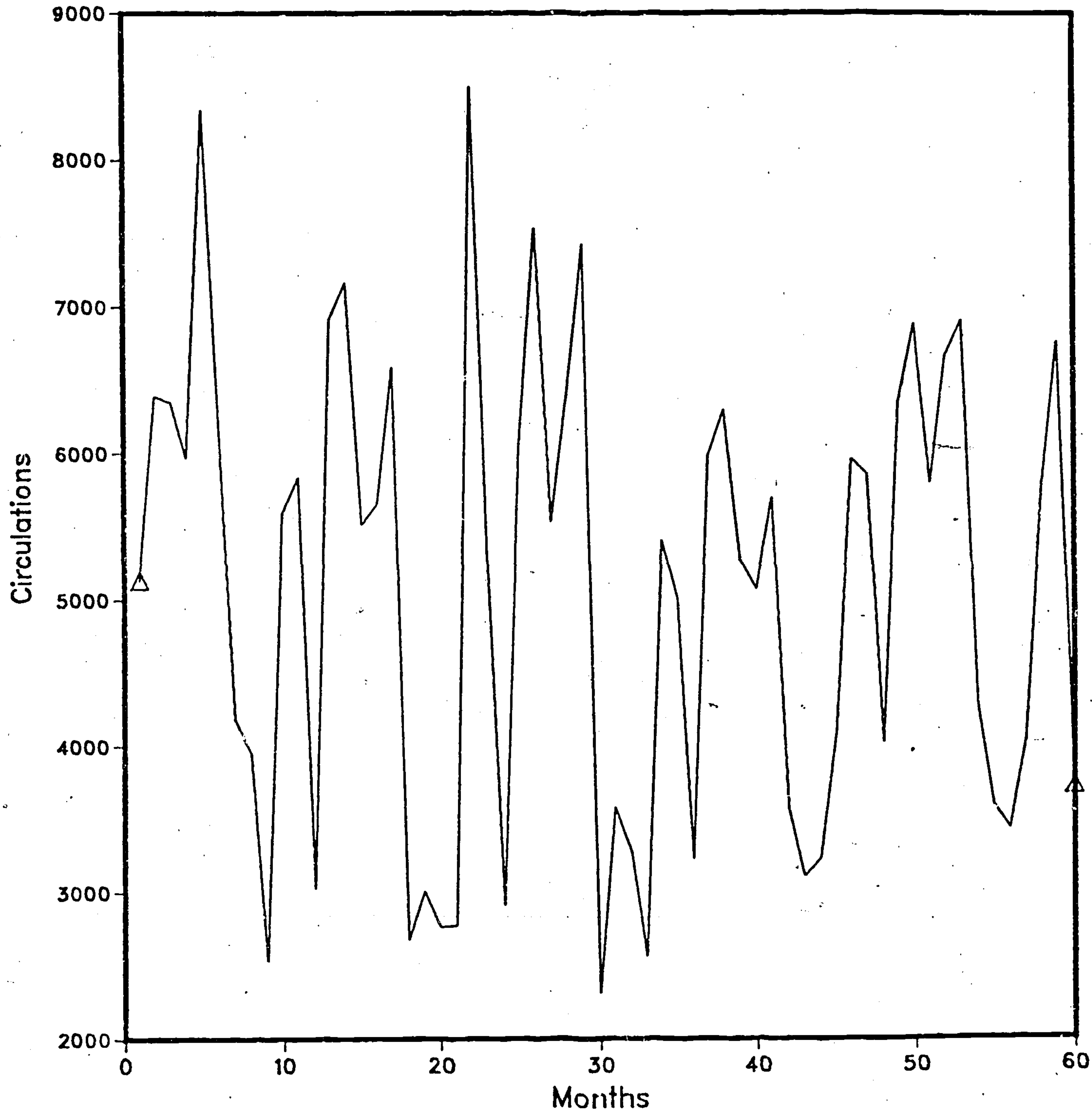
Library 11
Monthly Circulations



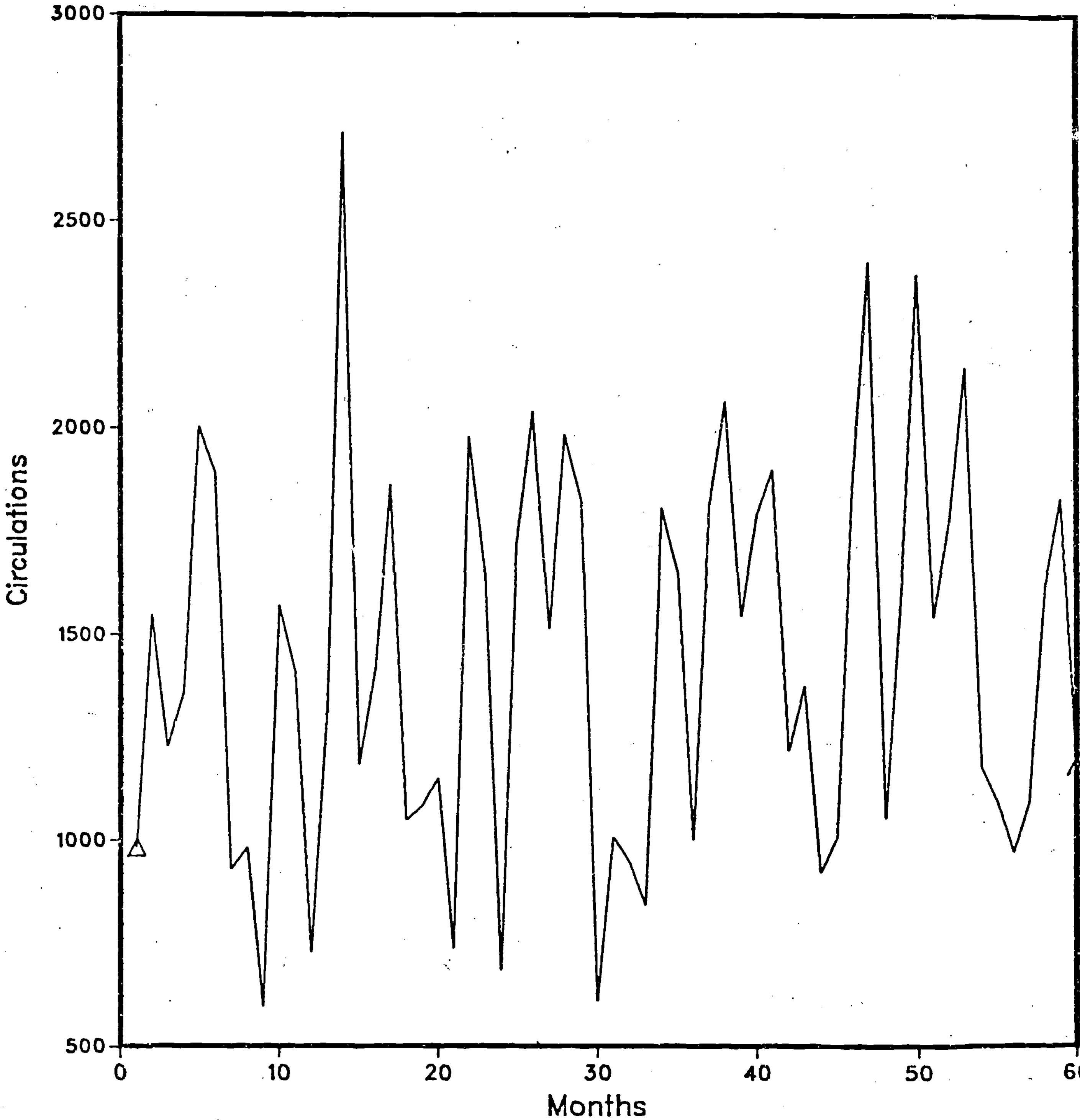
Library 12
Monthly Circulations

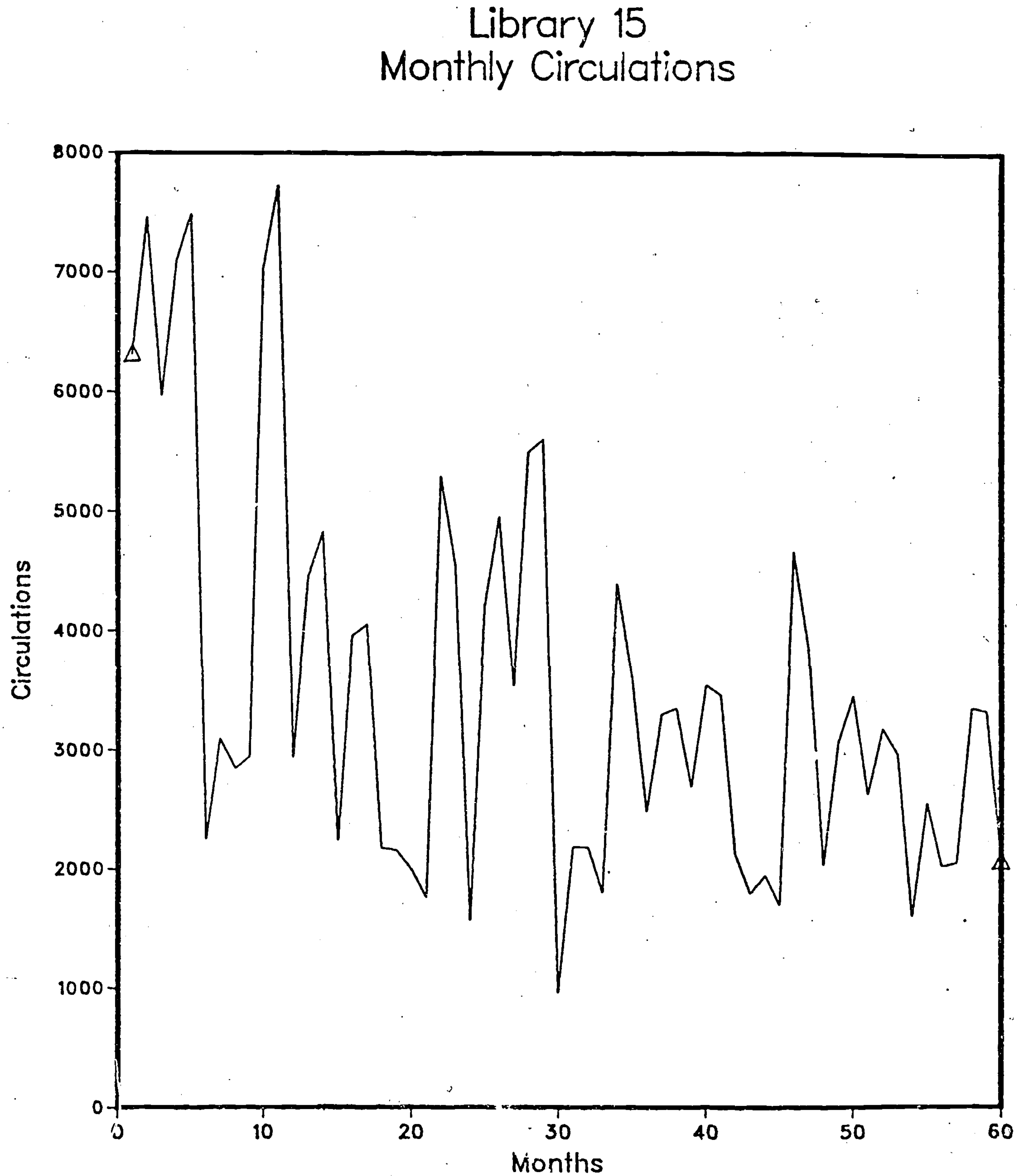


Library 13
Monthly Circulations

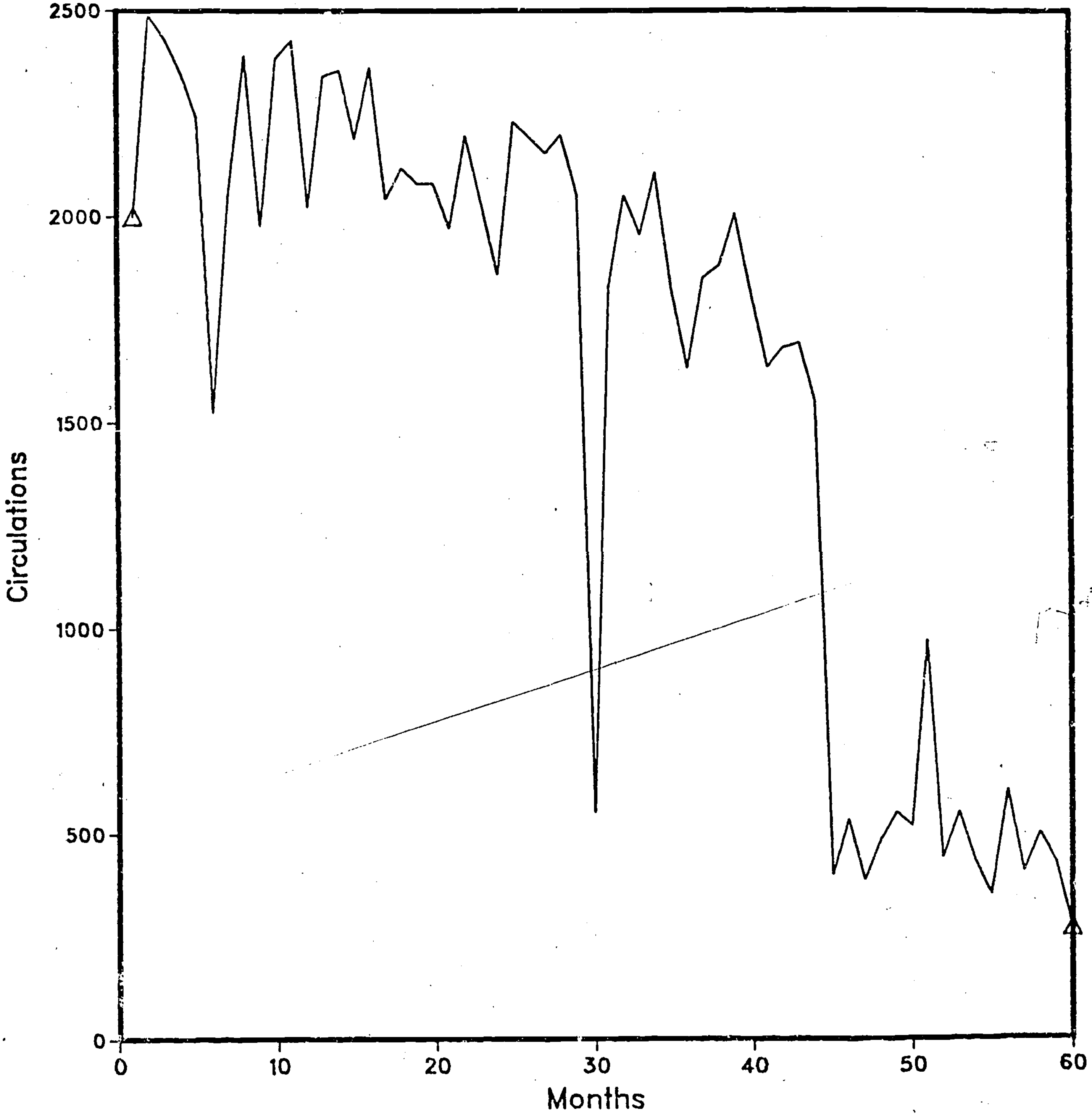


Library 14
Monthly Circulations

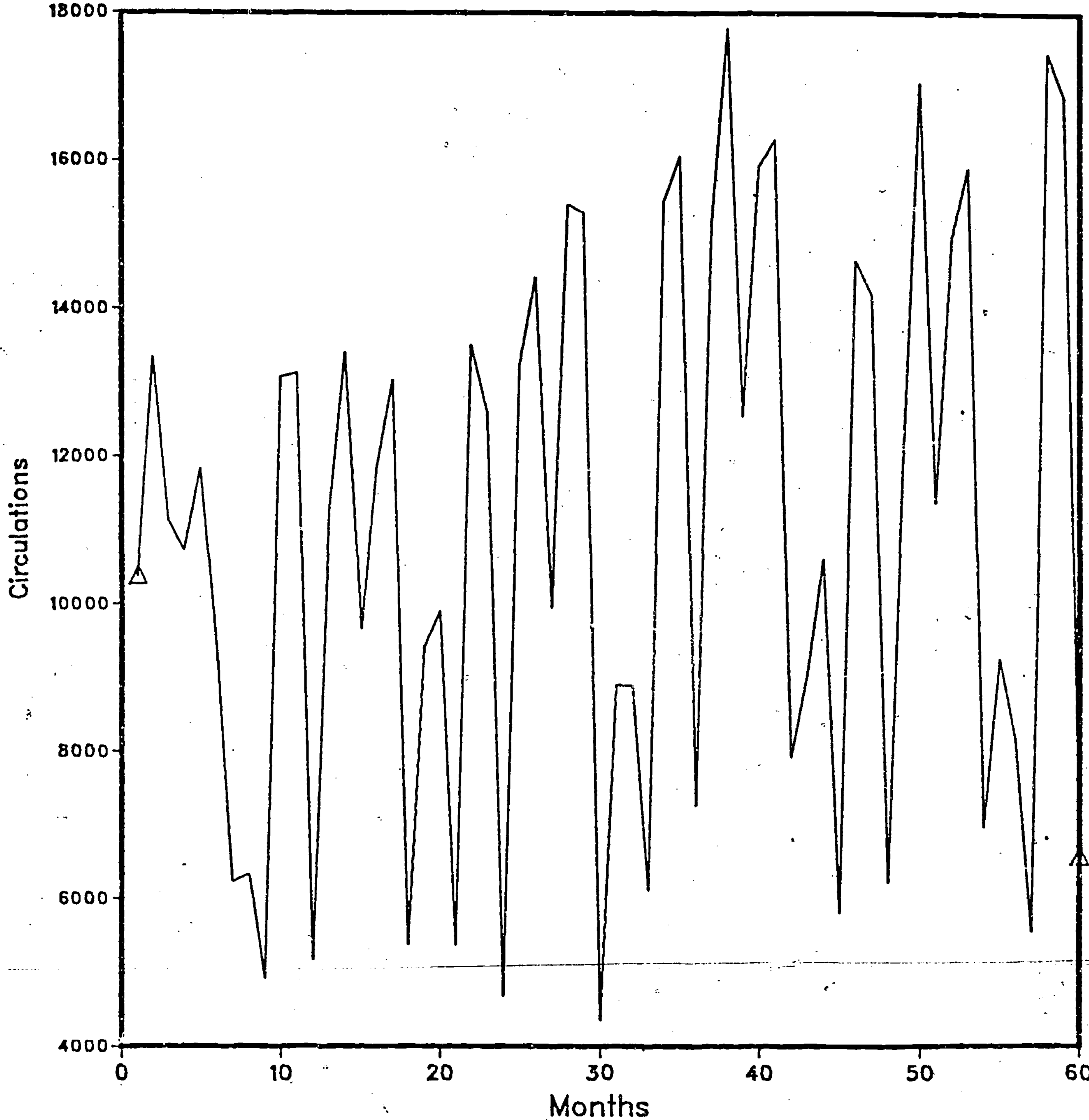



Library 15
Monthly Circulations
Circulations
Months
8000
7000
6000
5000
4000
3000
2000
1000
0
0
10
20
30
40
50
60

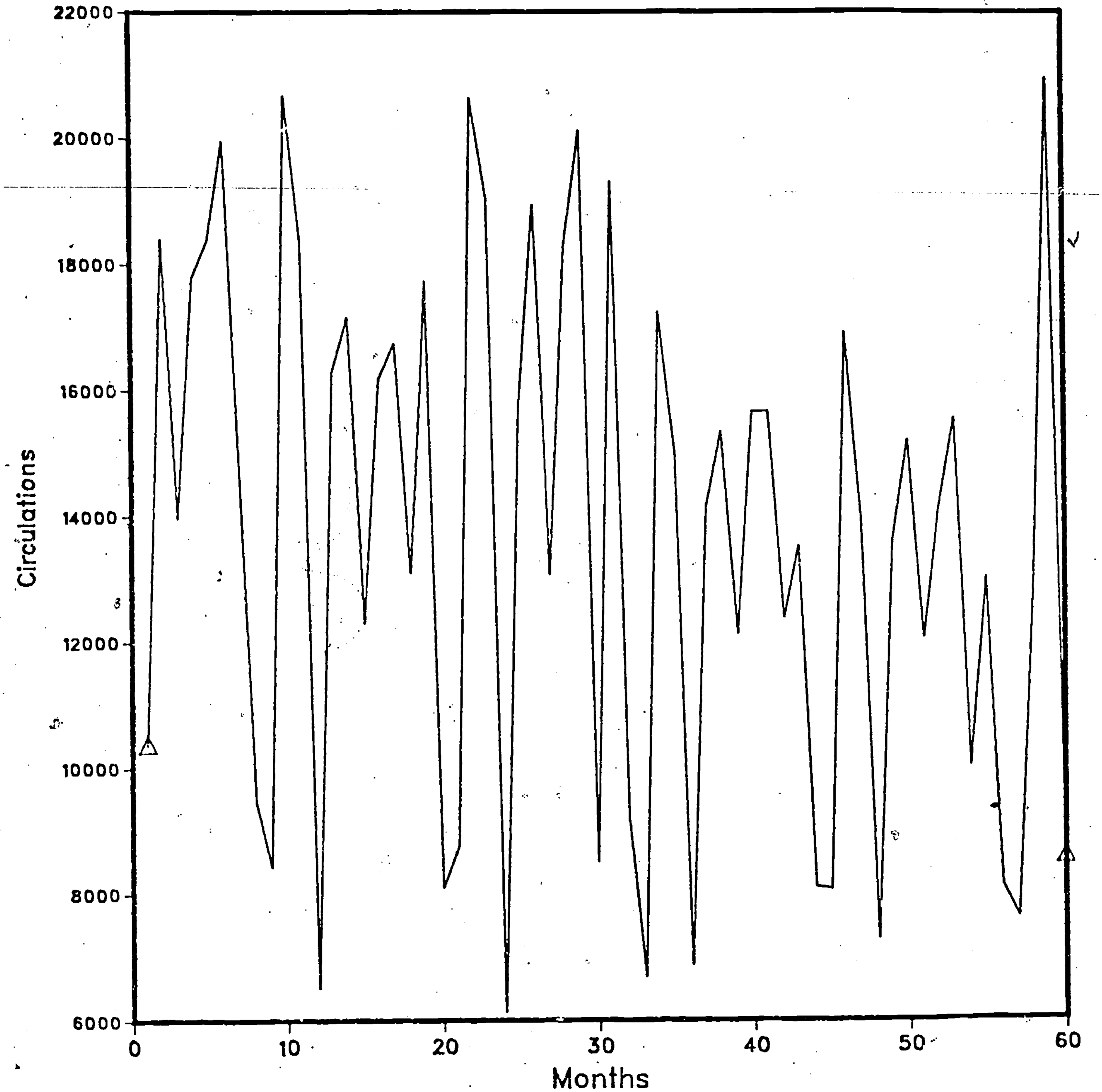
Library 16
Monthly Circulations

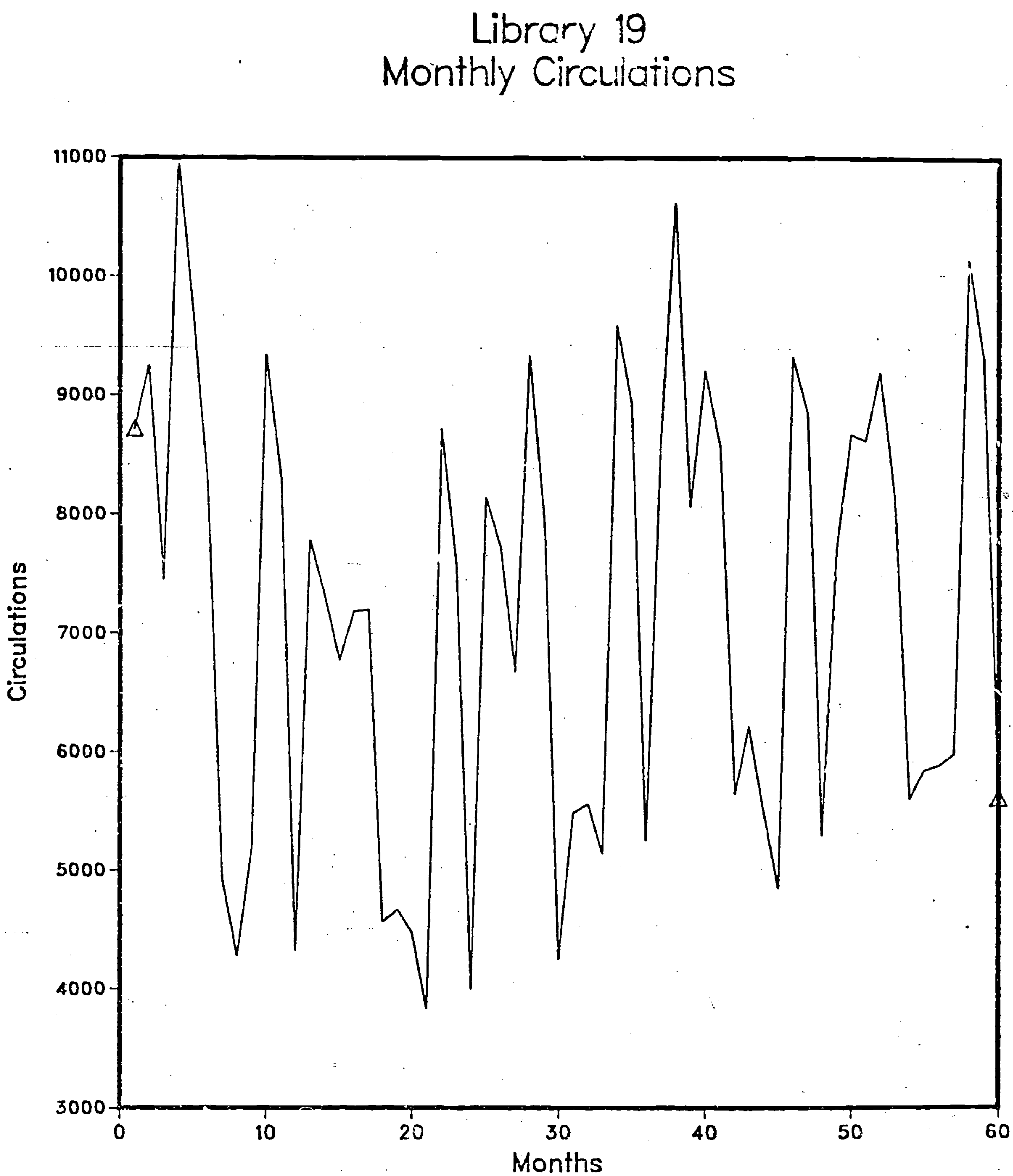


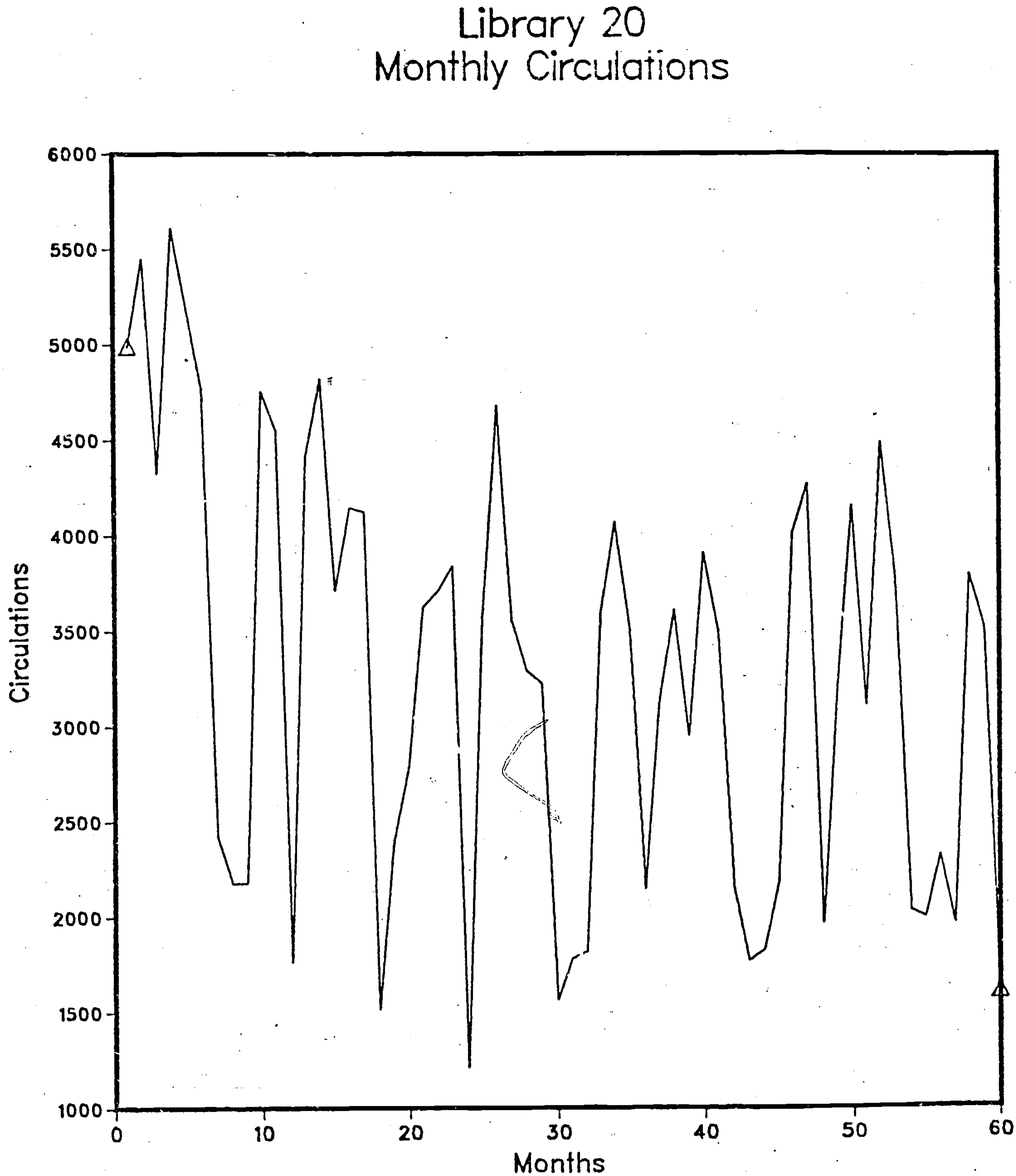
Library 17
Monthly Circulations



Library 18
Monthly Circulations

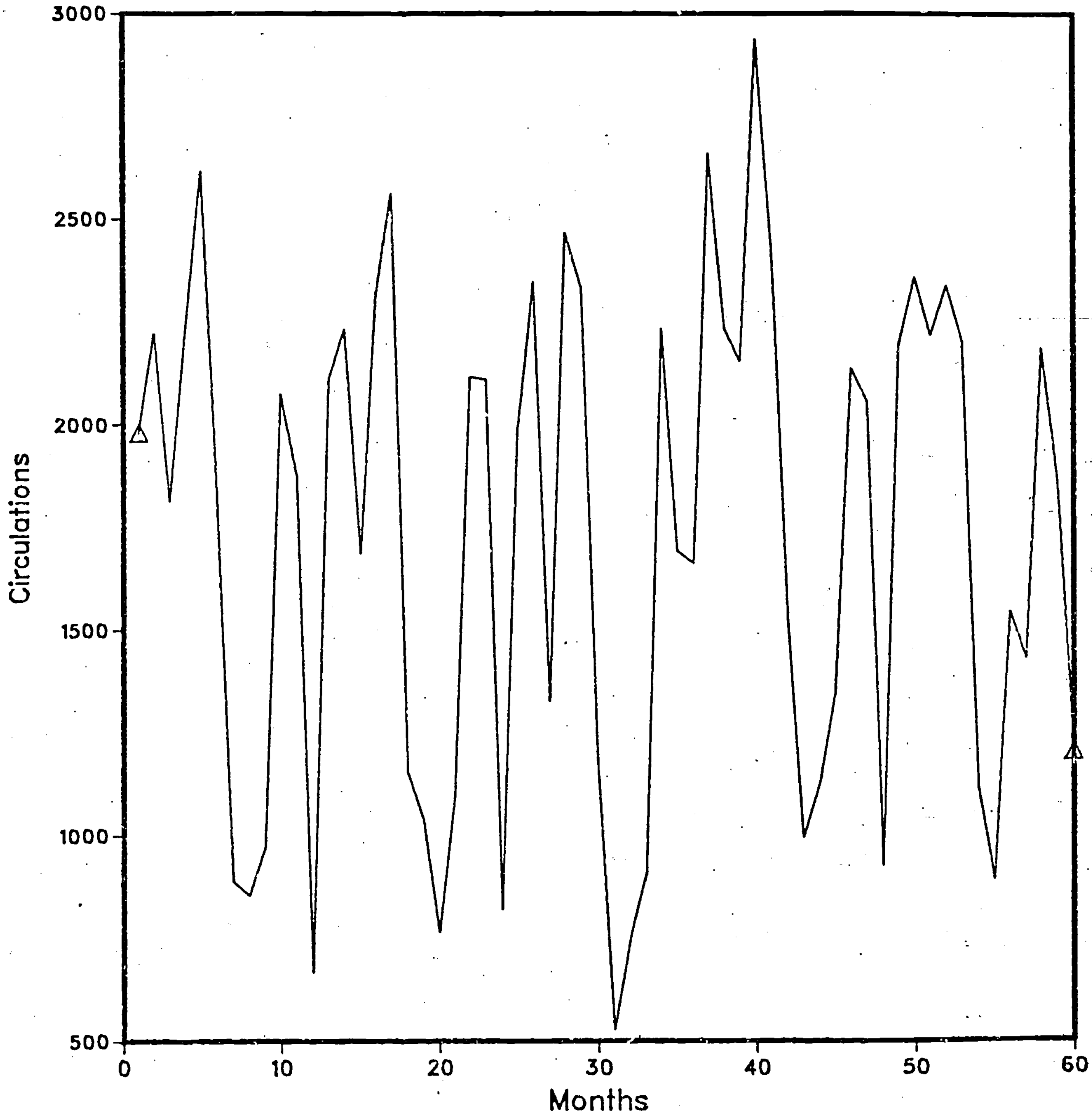



Library 19
Monthly Circulations
11000
10000
9000
8000
7000
6000
5000
4000
3000
Circulations
0
10
20
30
40
50
60
Months

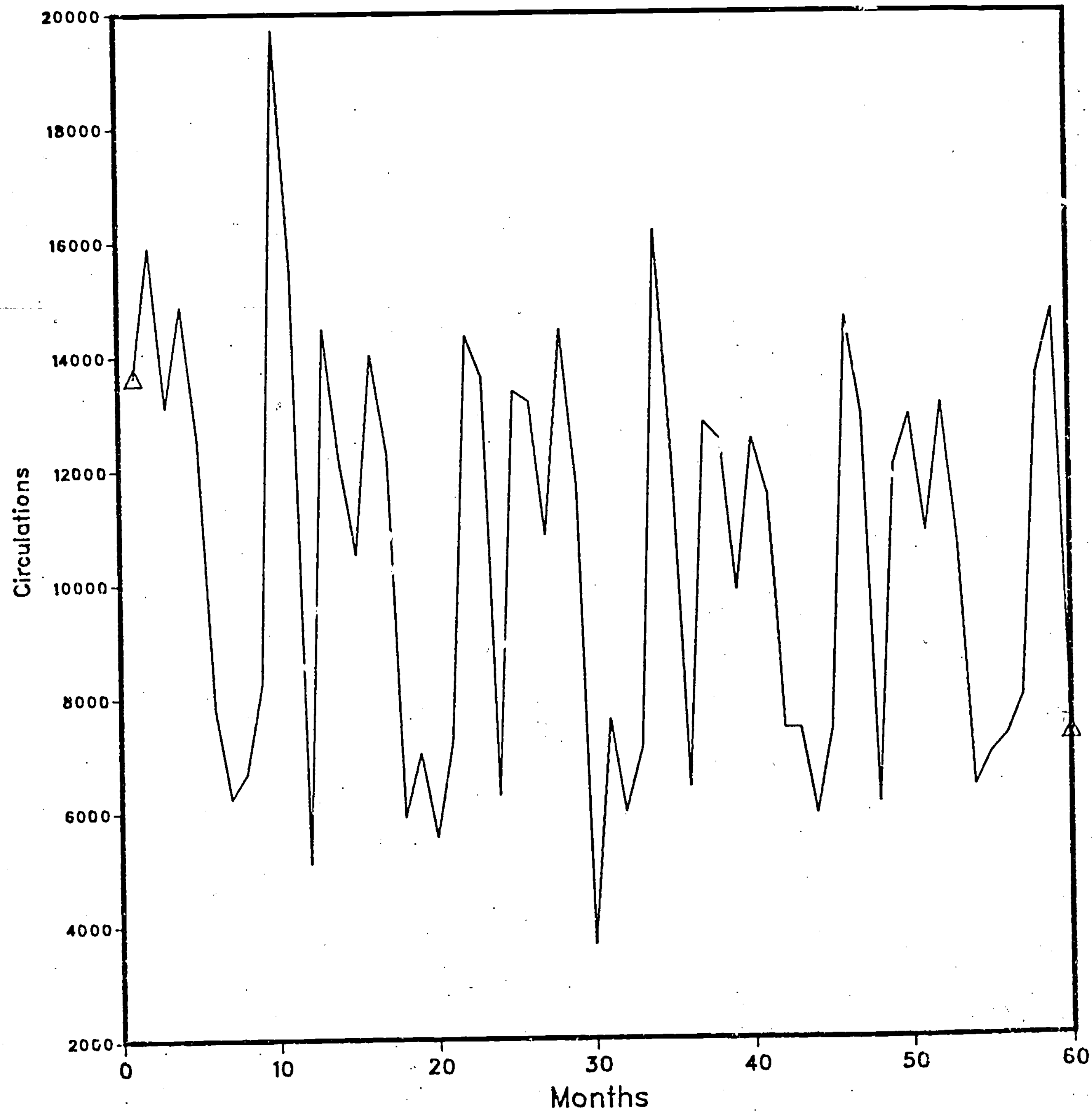


Library 20
Monthly Circulations
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
Circulations
0
10
20
30
40
50
60
Months

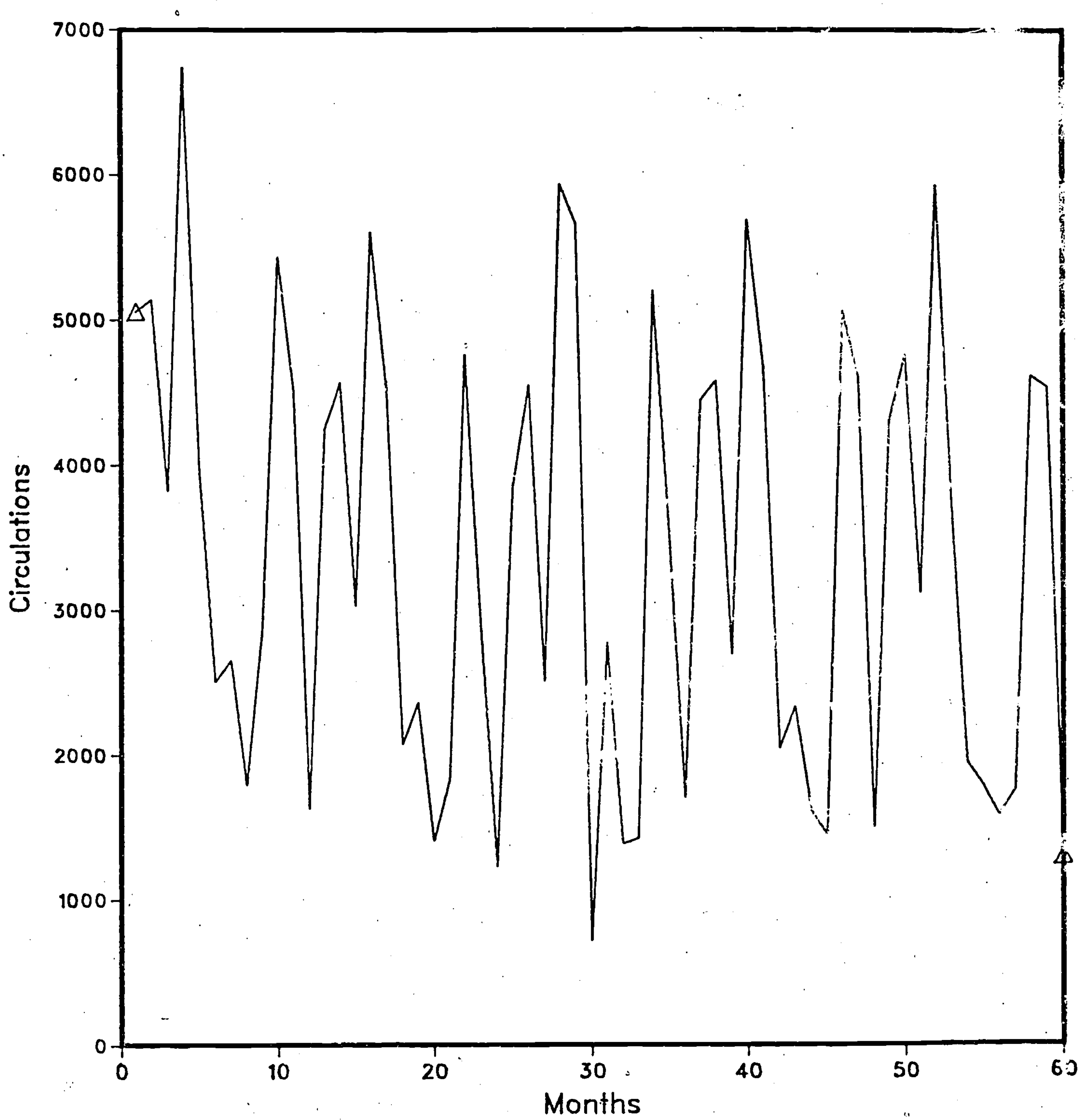
Library 21
Monthly Circulations



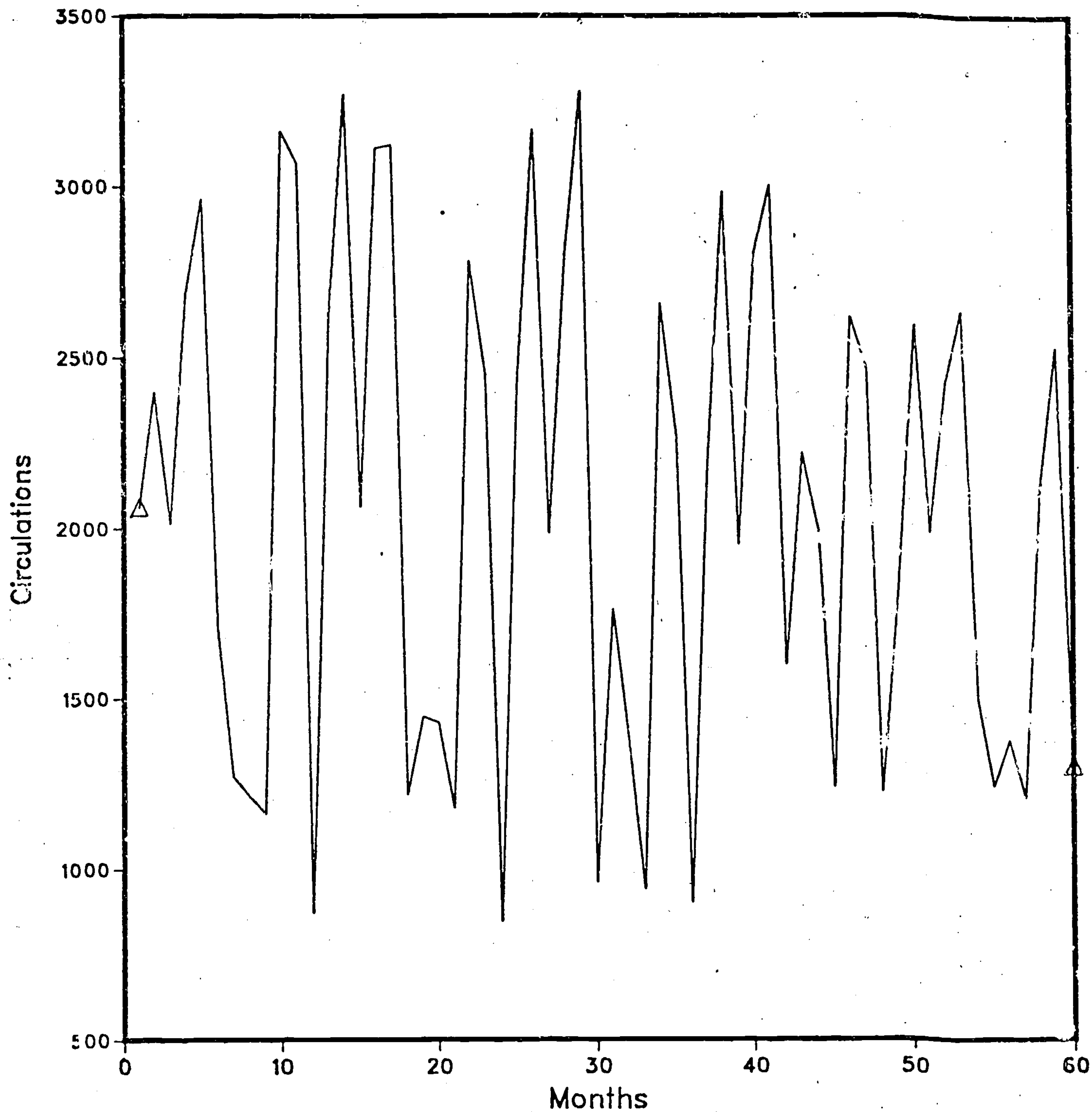
Library 22
Monthly Circulations



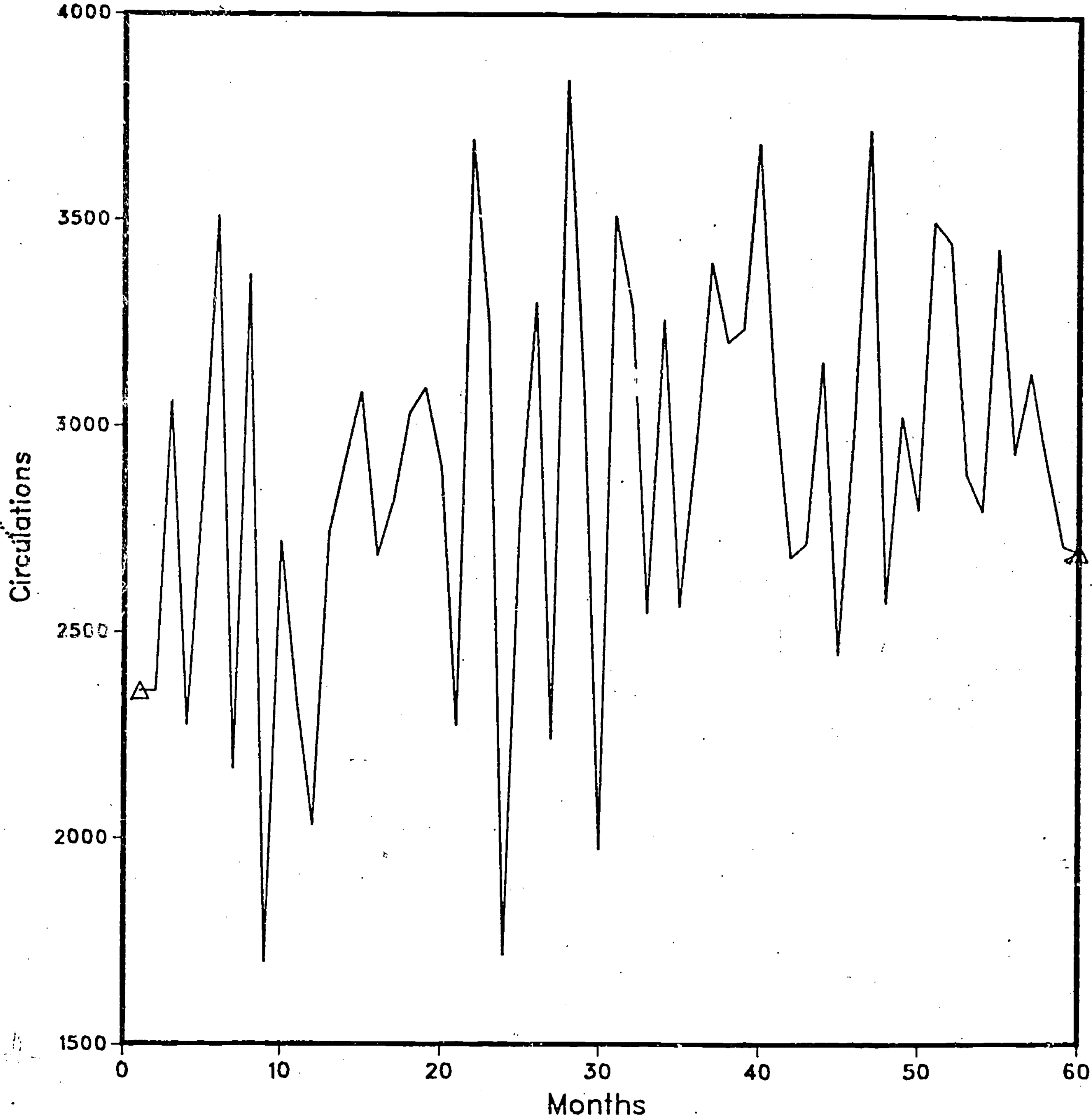
Library 23
Monthly Circulations



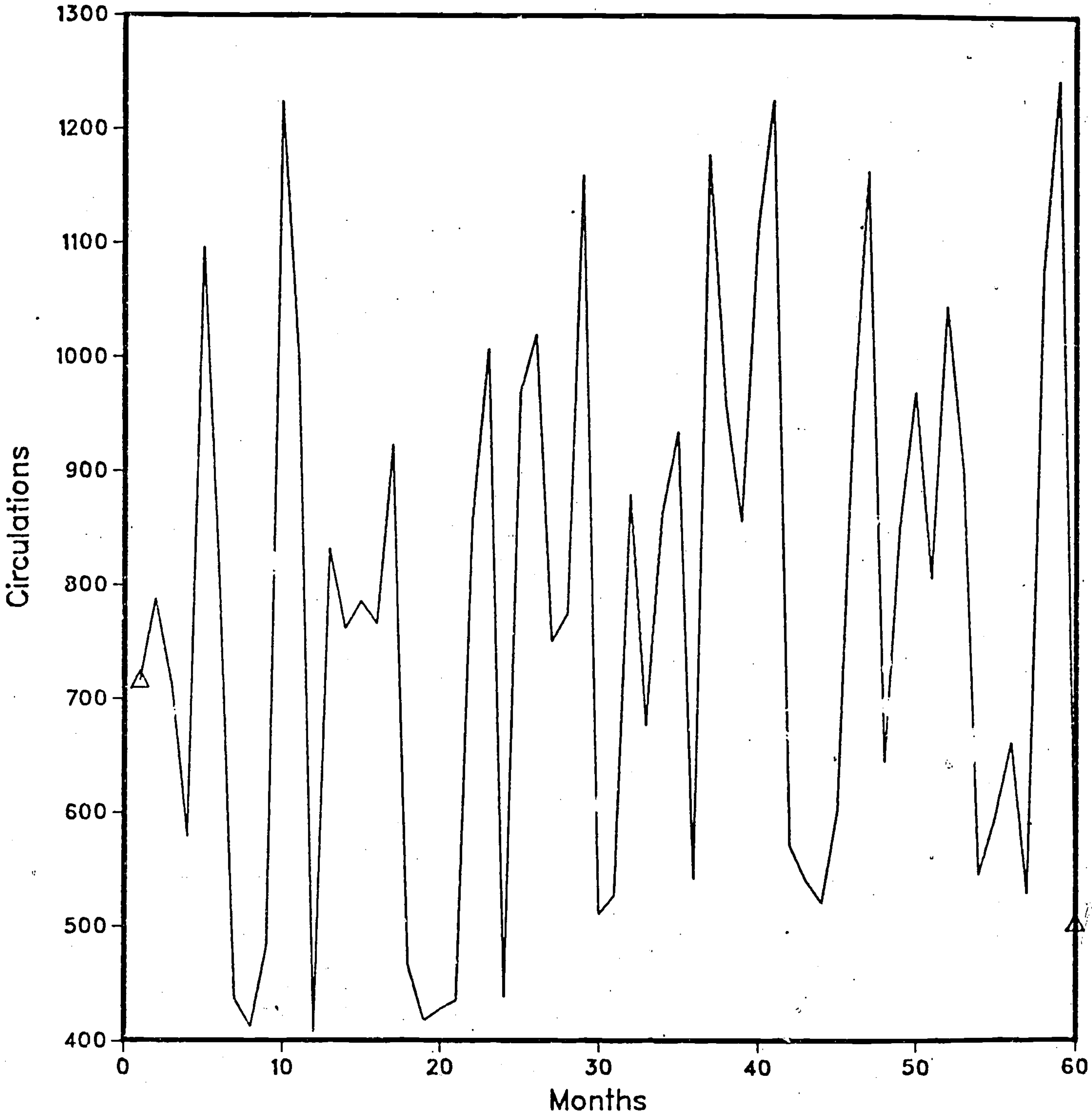
Library 24
Monthly Circulations

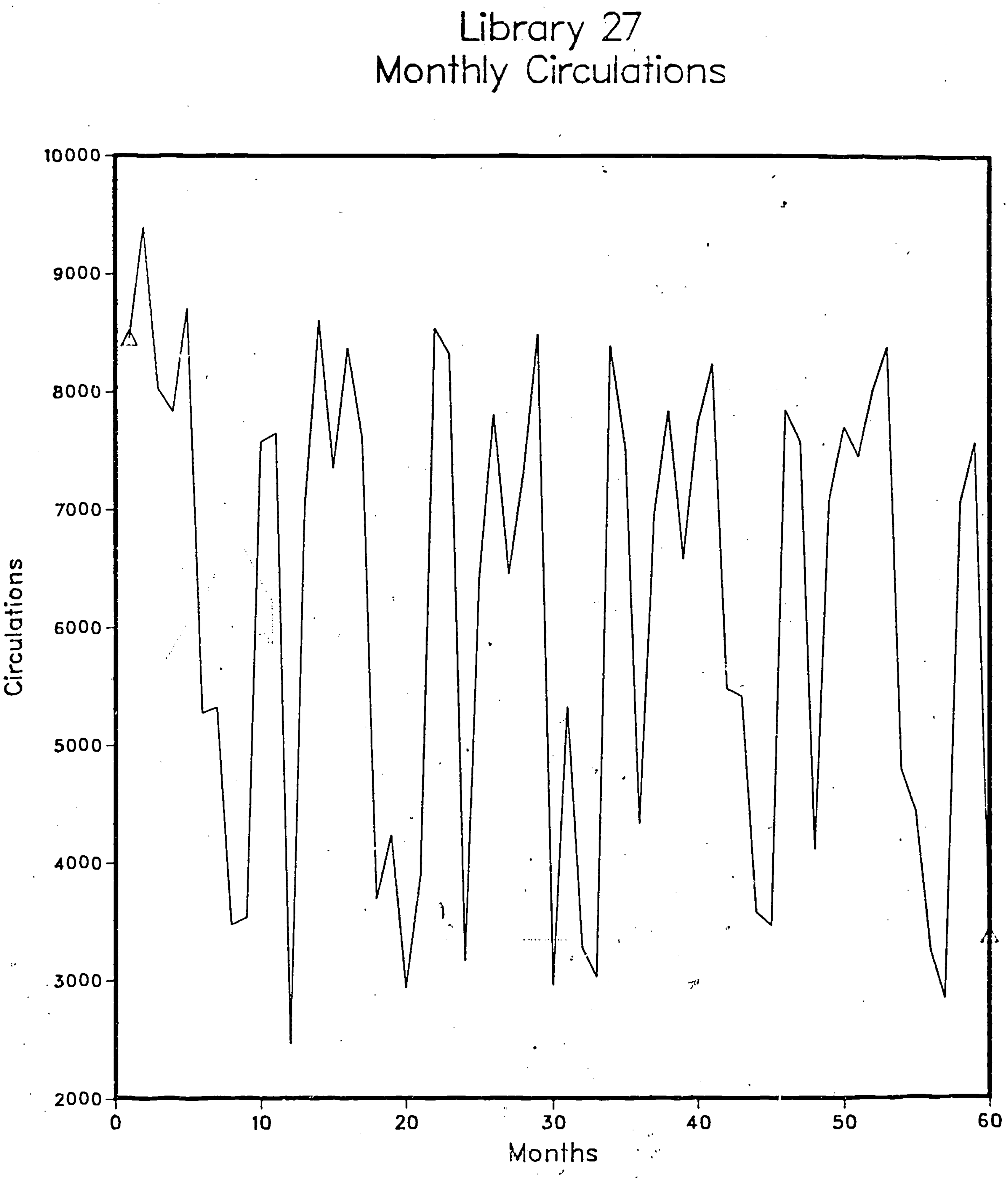


Library 25
Monthly Circulations



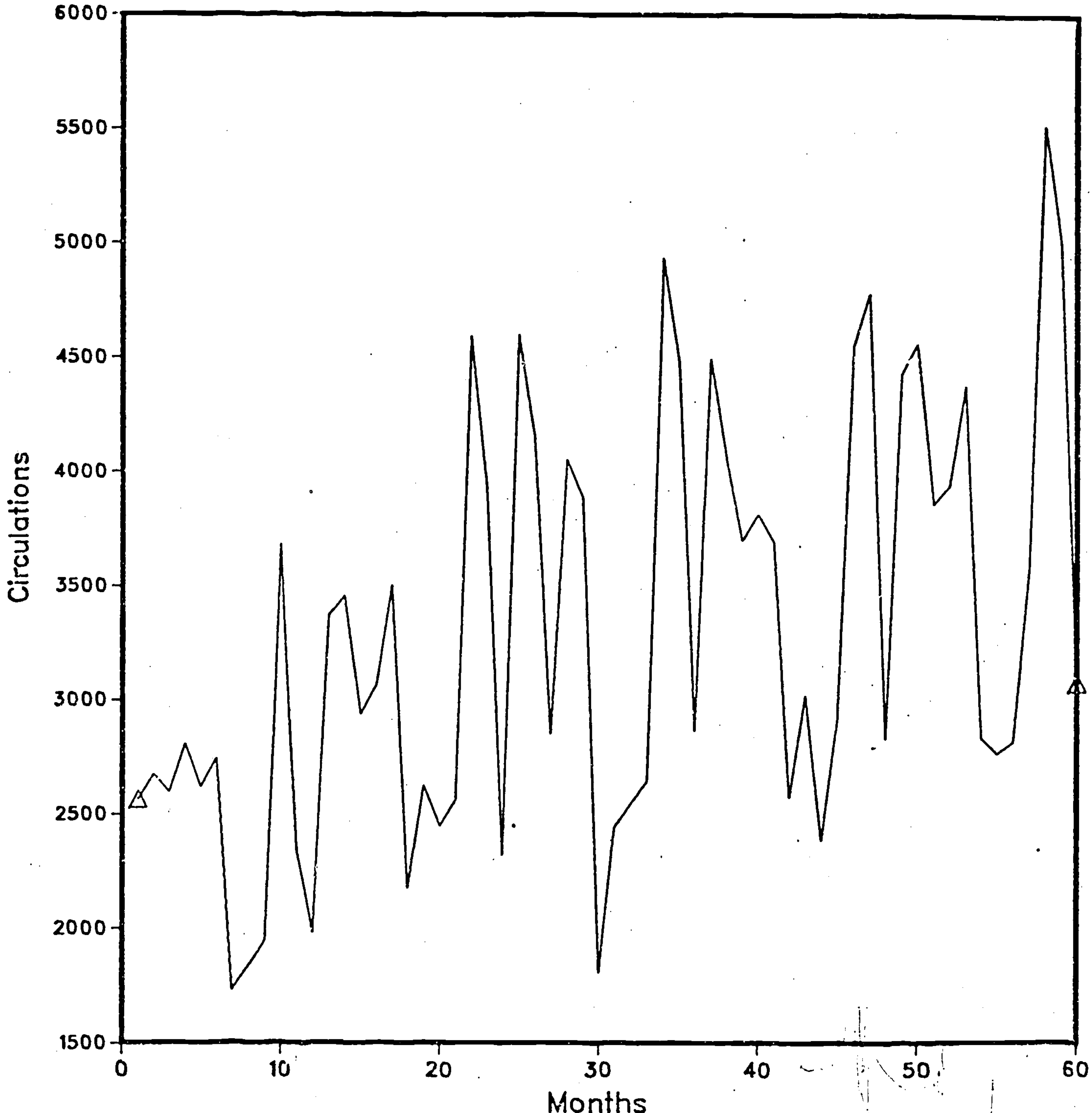
Library 26
Monthly Circulations

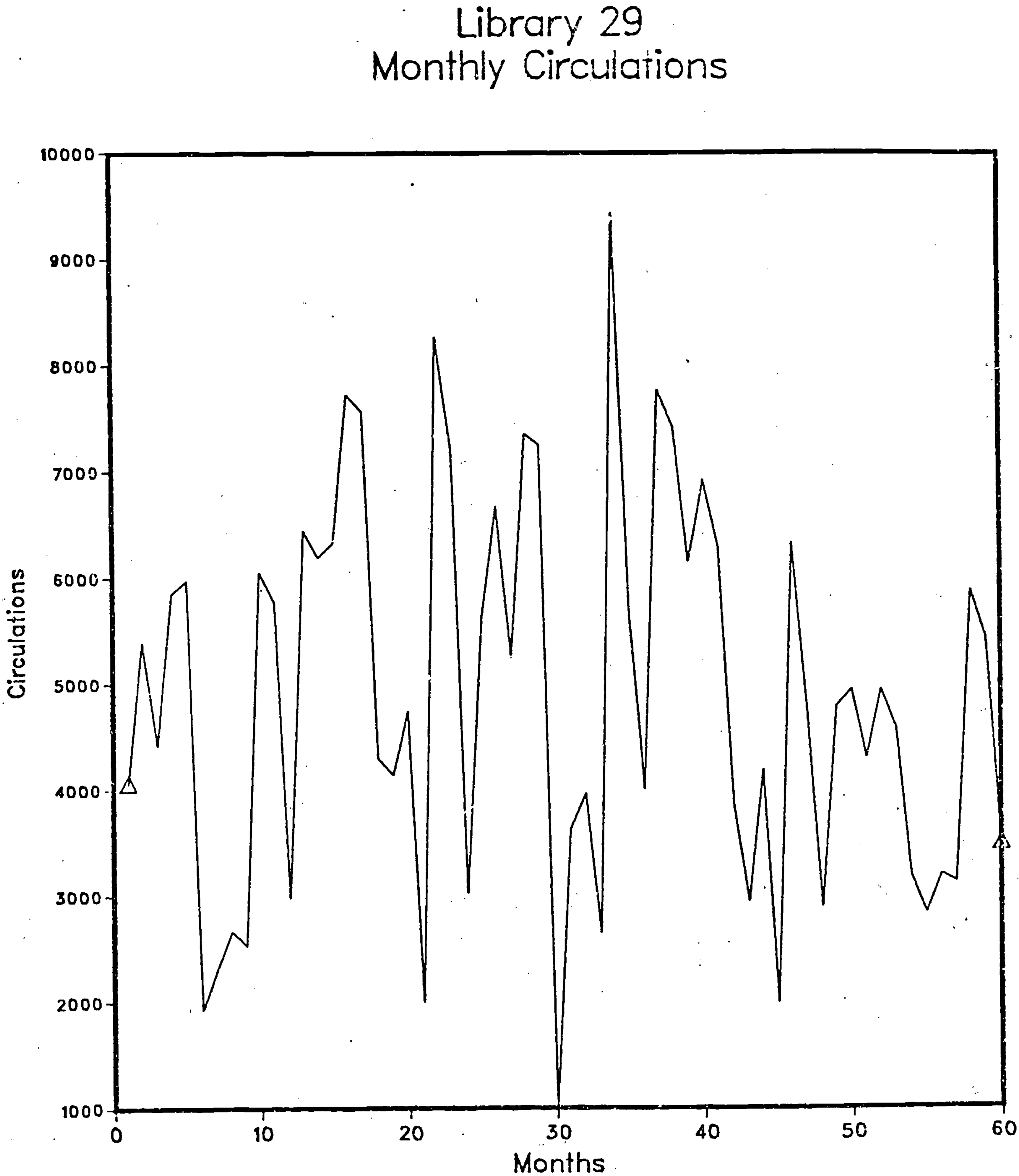




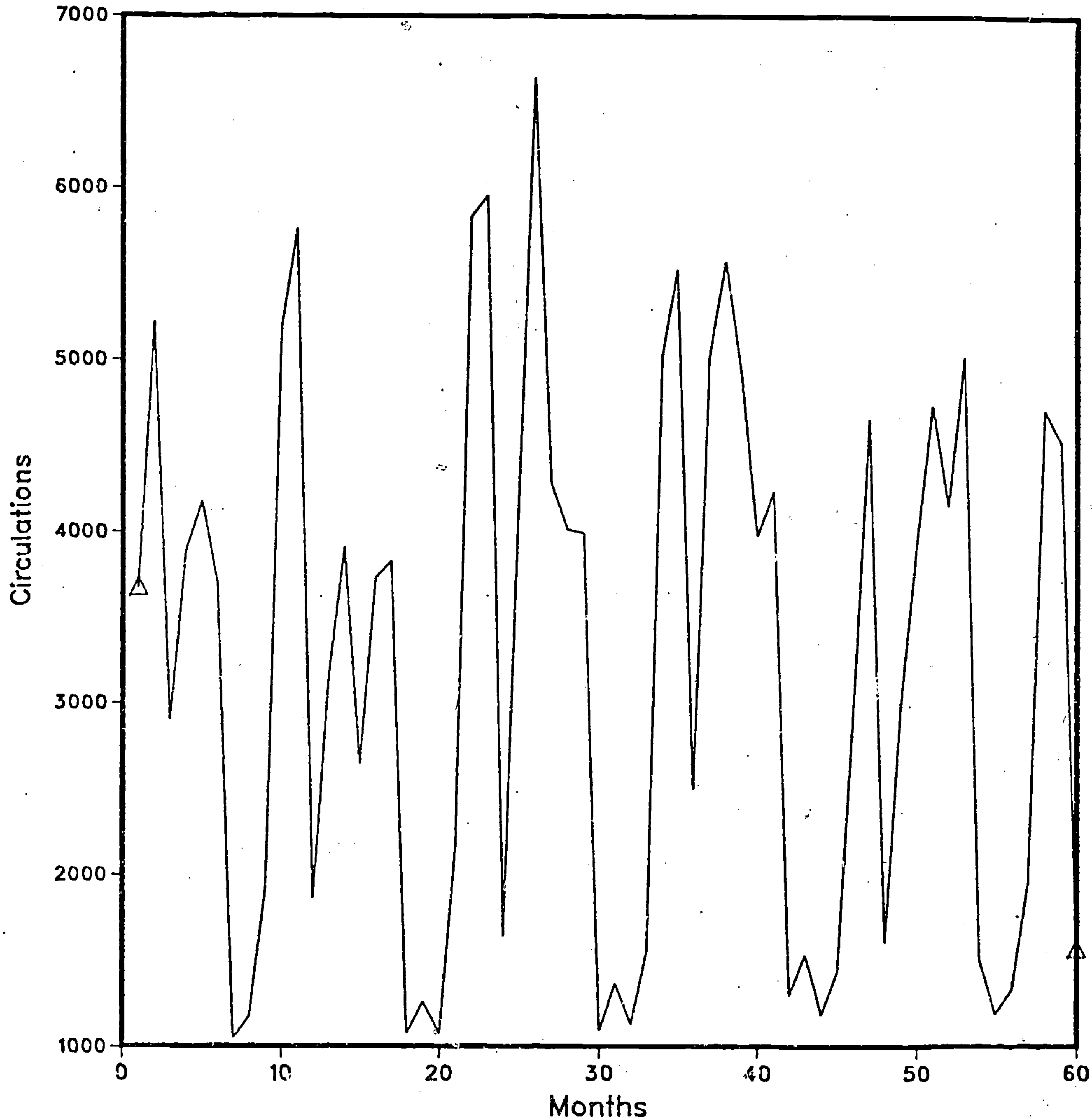
Library 27
Monthly Circulations
10000
9000
8000
7000
6000
5000
4000
3000
2000
Circulations
0
10
20
30
40
50
60
Months

Library 28
Monthly Circulations

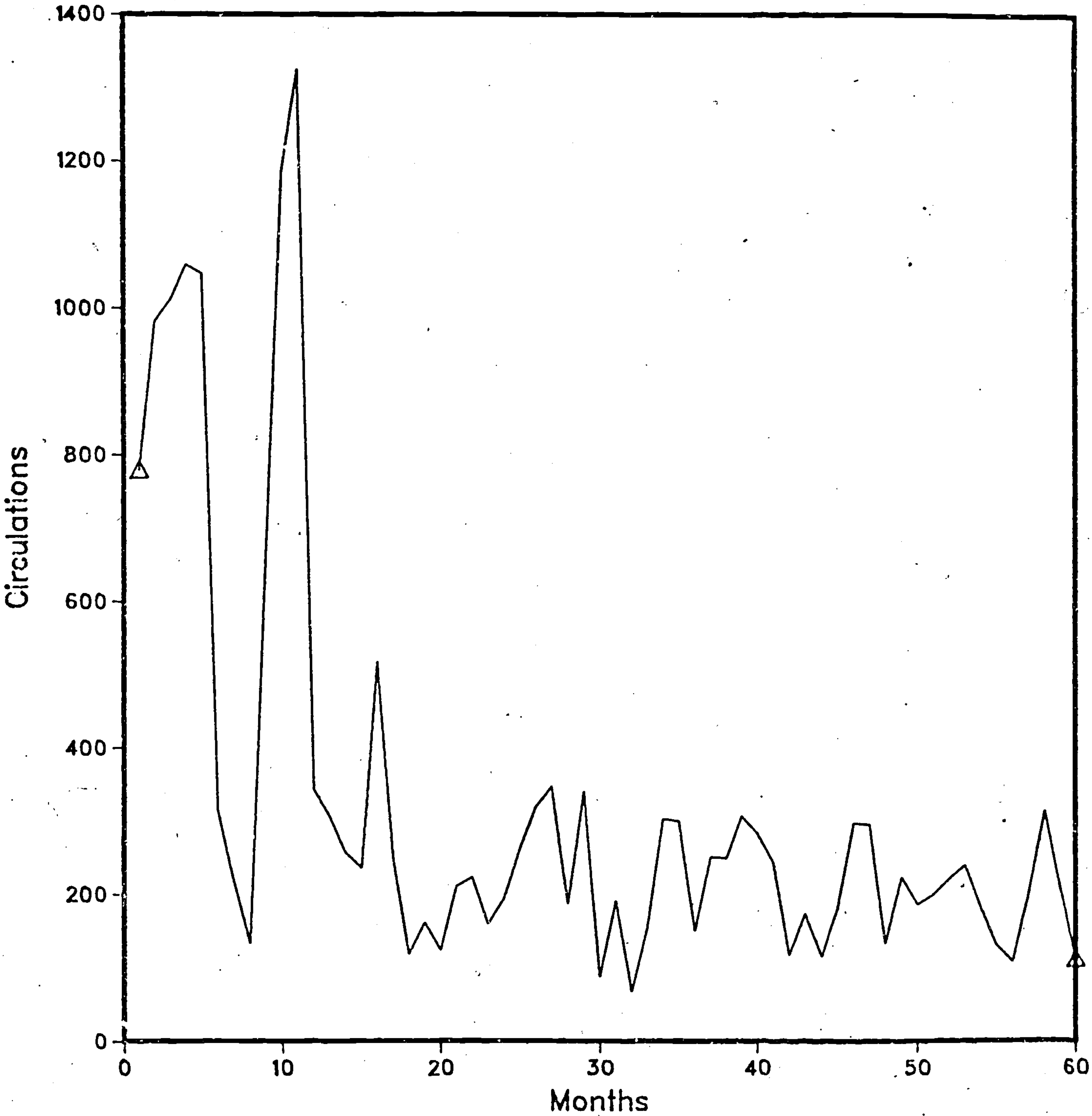



Library 29
Monthly Circulations
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
Circulations
0
10
20
30
40
50
60
Months

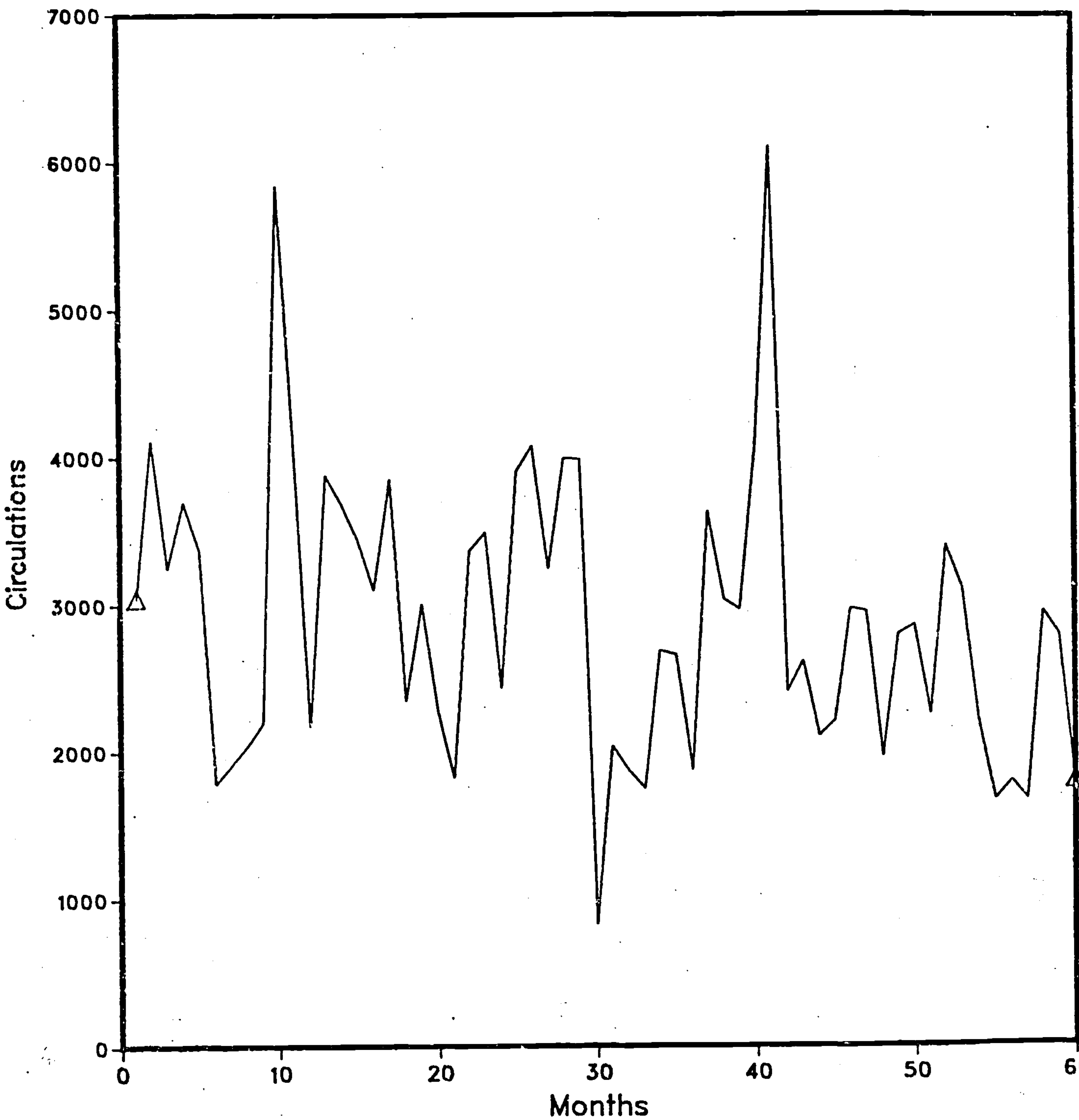
Library 30
Monthly Circulations

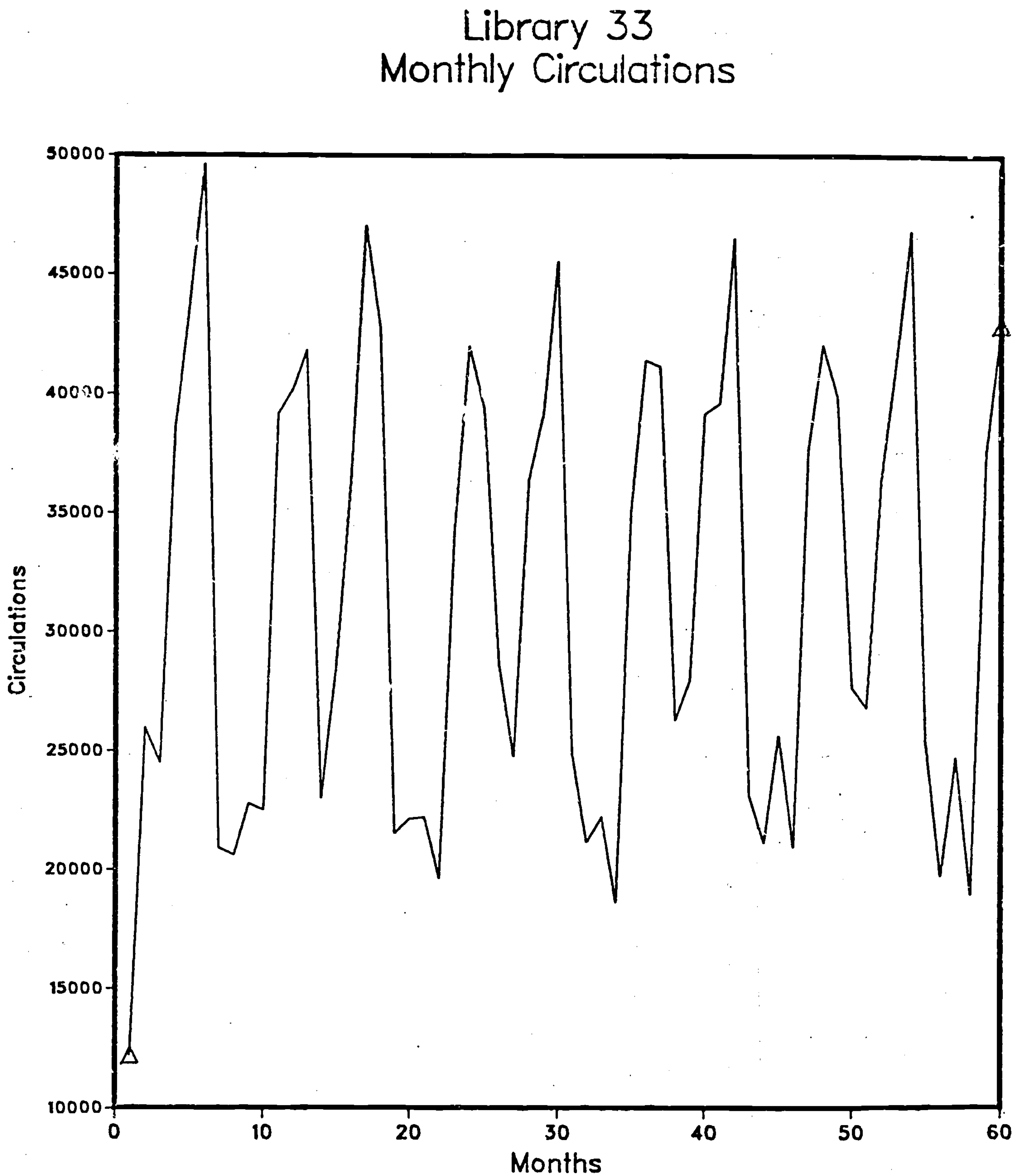


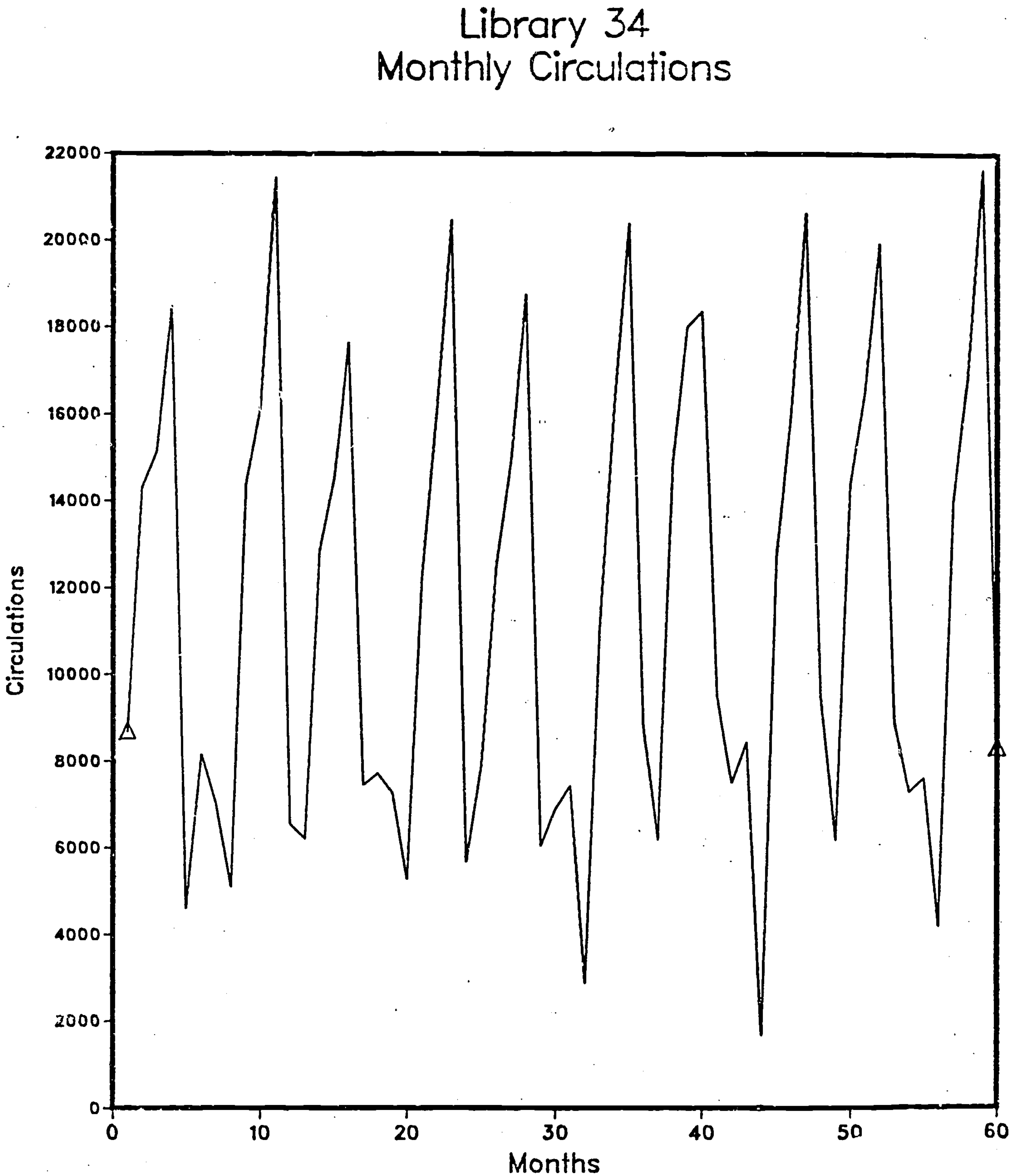
Library 31
Monthly Circulations



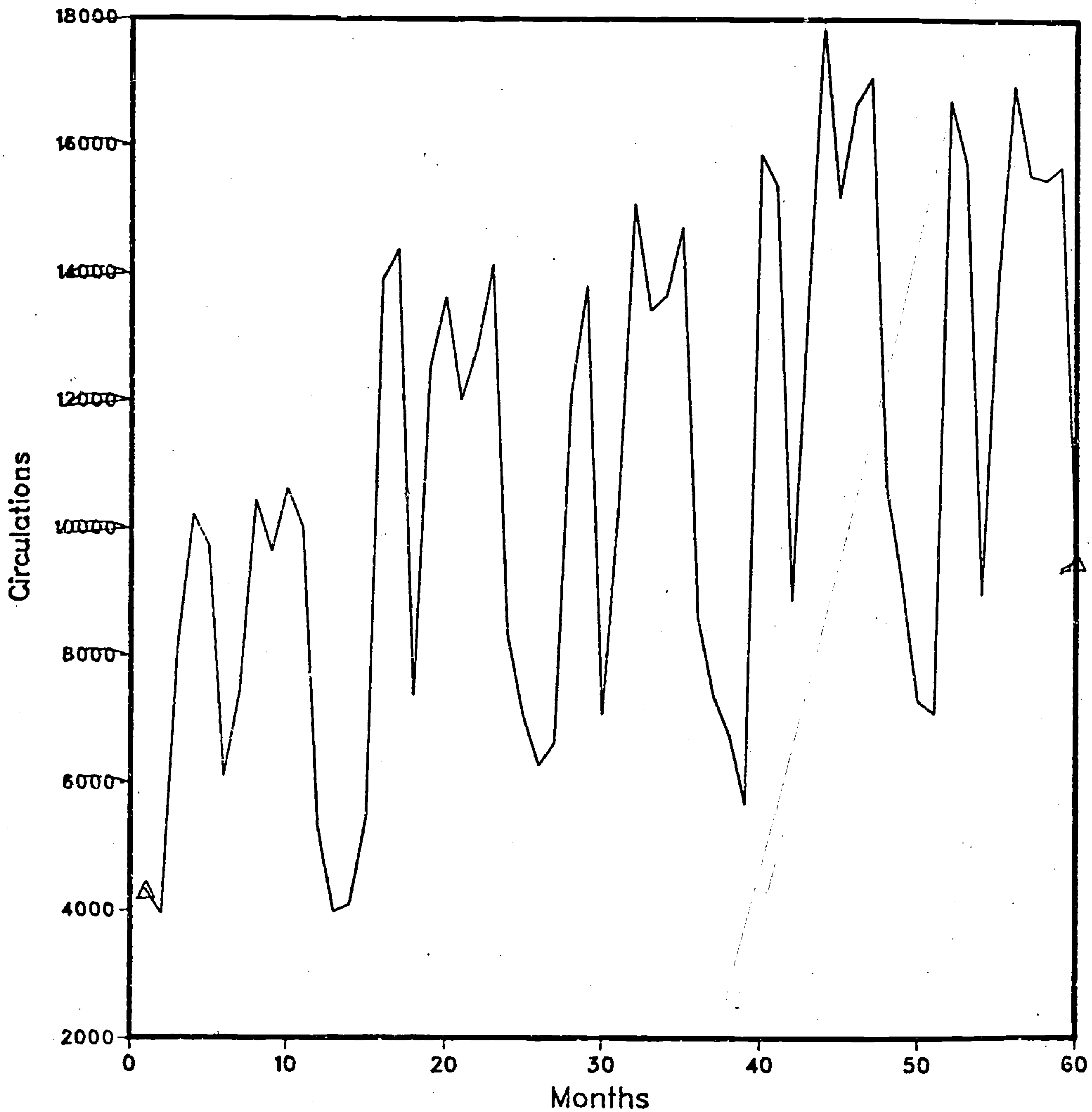
Library 32
Monthly Circulations



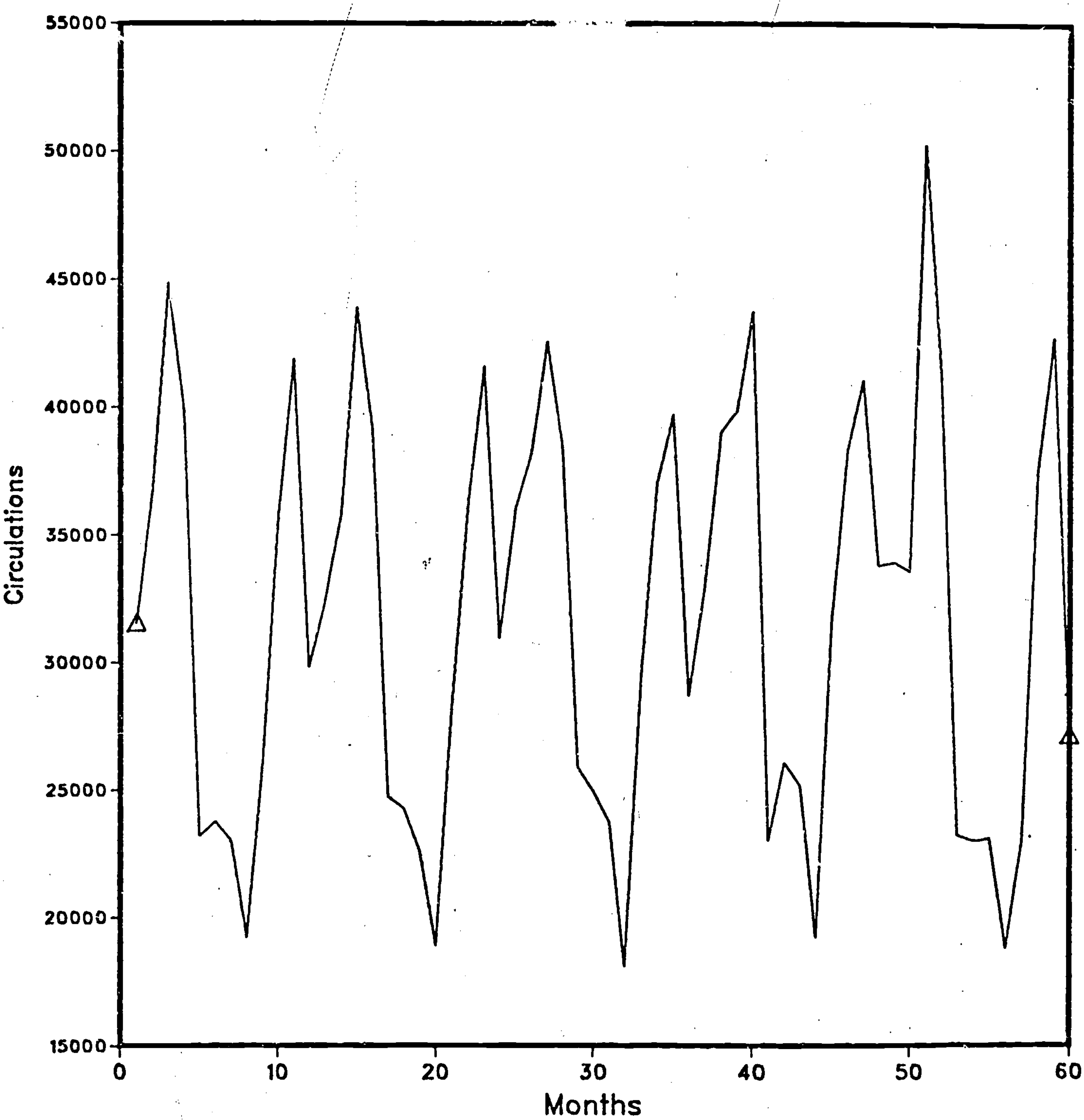

Library 33
Monthly Circulations
50000
45000
40000
35000
30000
25000
20000
15000
10000
Circulations
0
10
20
30
40
50
60
Months


Library 34
Monthly Circulations
Circulations
Months
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
20000
22000
0
10
20
30
40
50
60

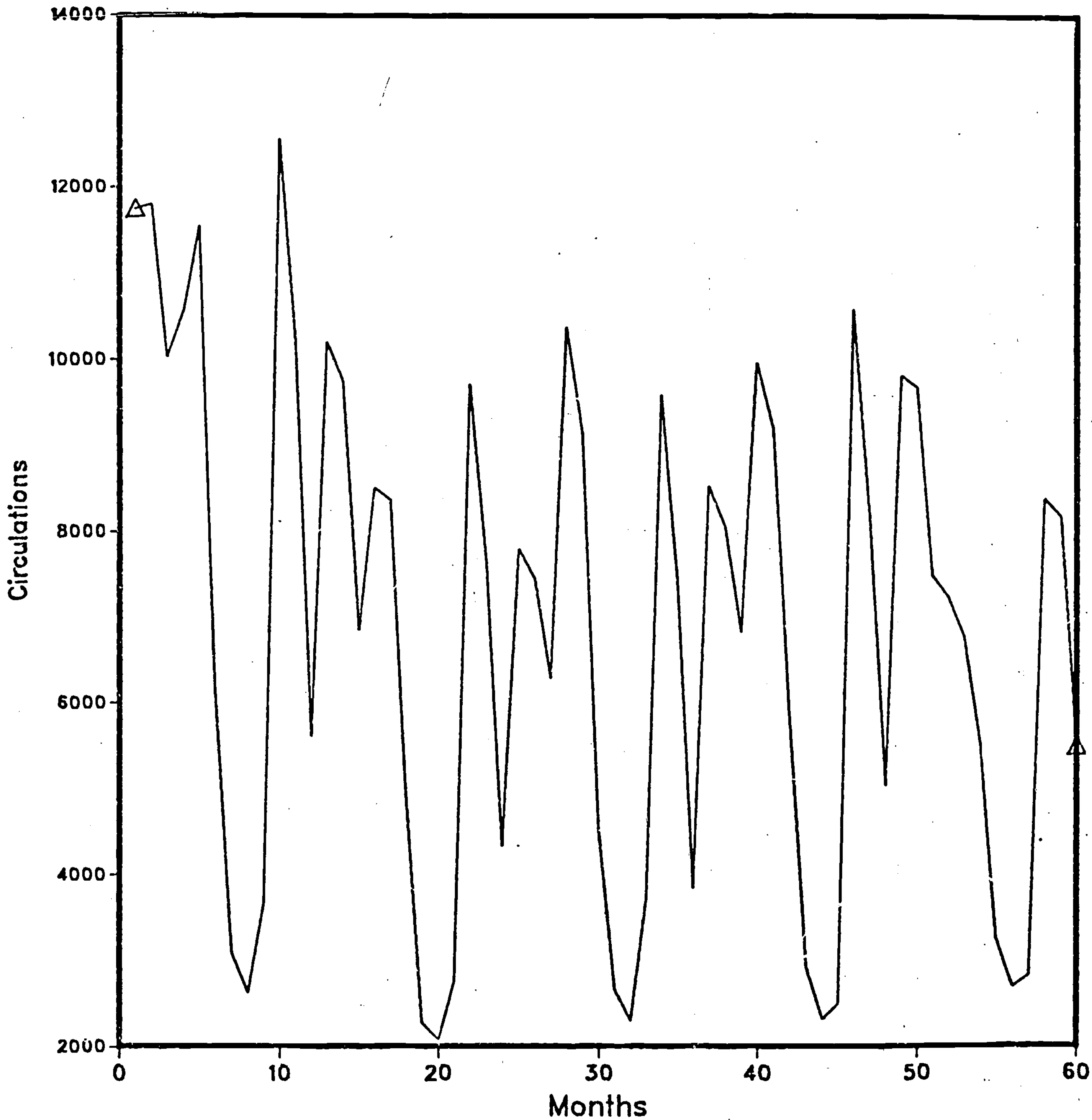
Library 35
Monthly Circulations

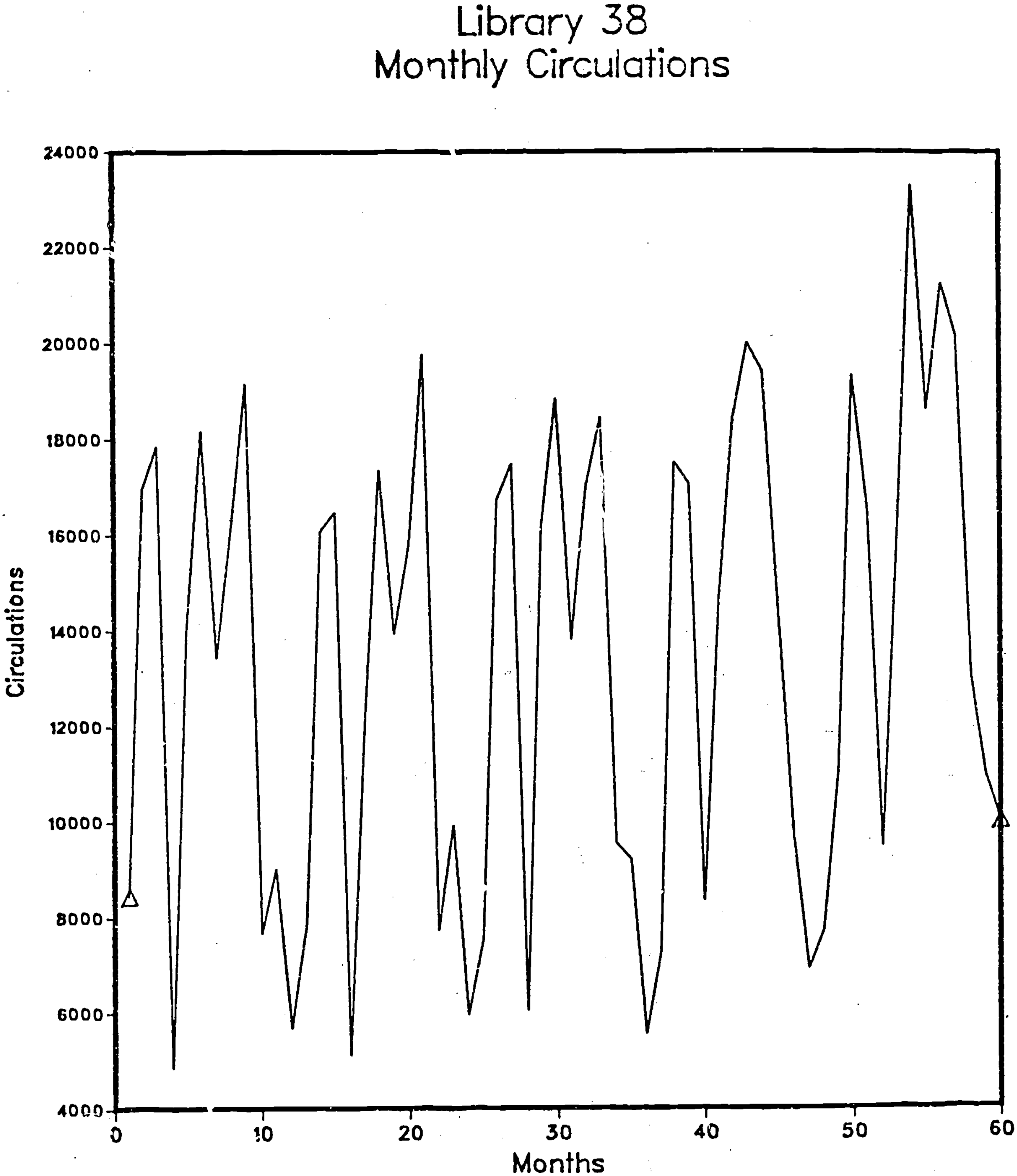


Library 36
Monthly Circulations



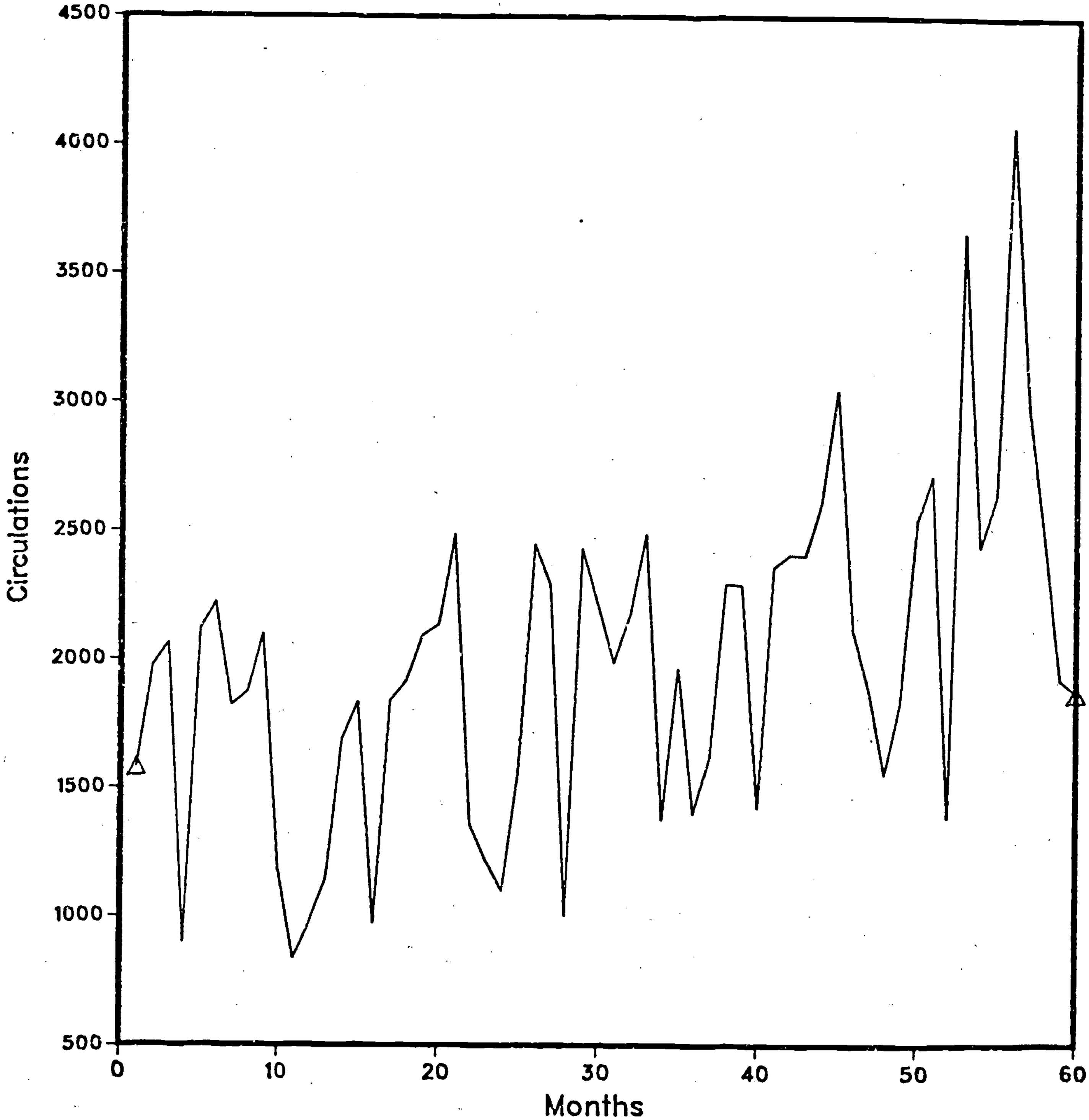
Library 37
Monthly Circulations



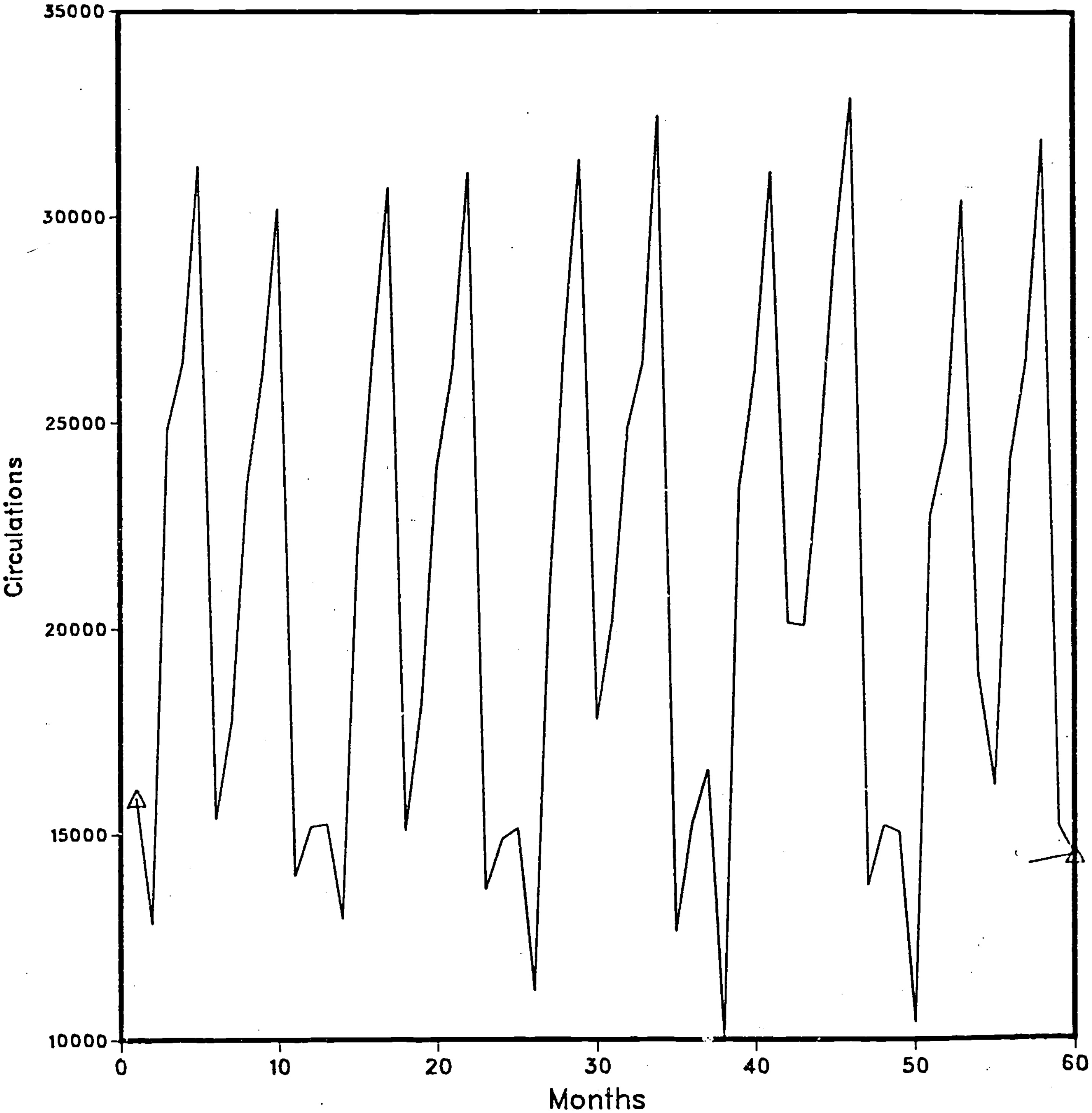


Library 38
Monthly Circulations
Circulations
24000
22000
20000
18000
16000
14000
12000
10000
8000
6000
4000
0
10
20
30
40
50
60
Months

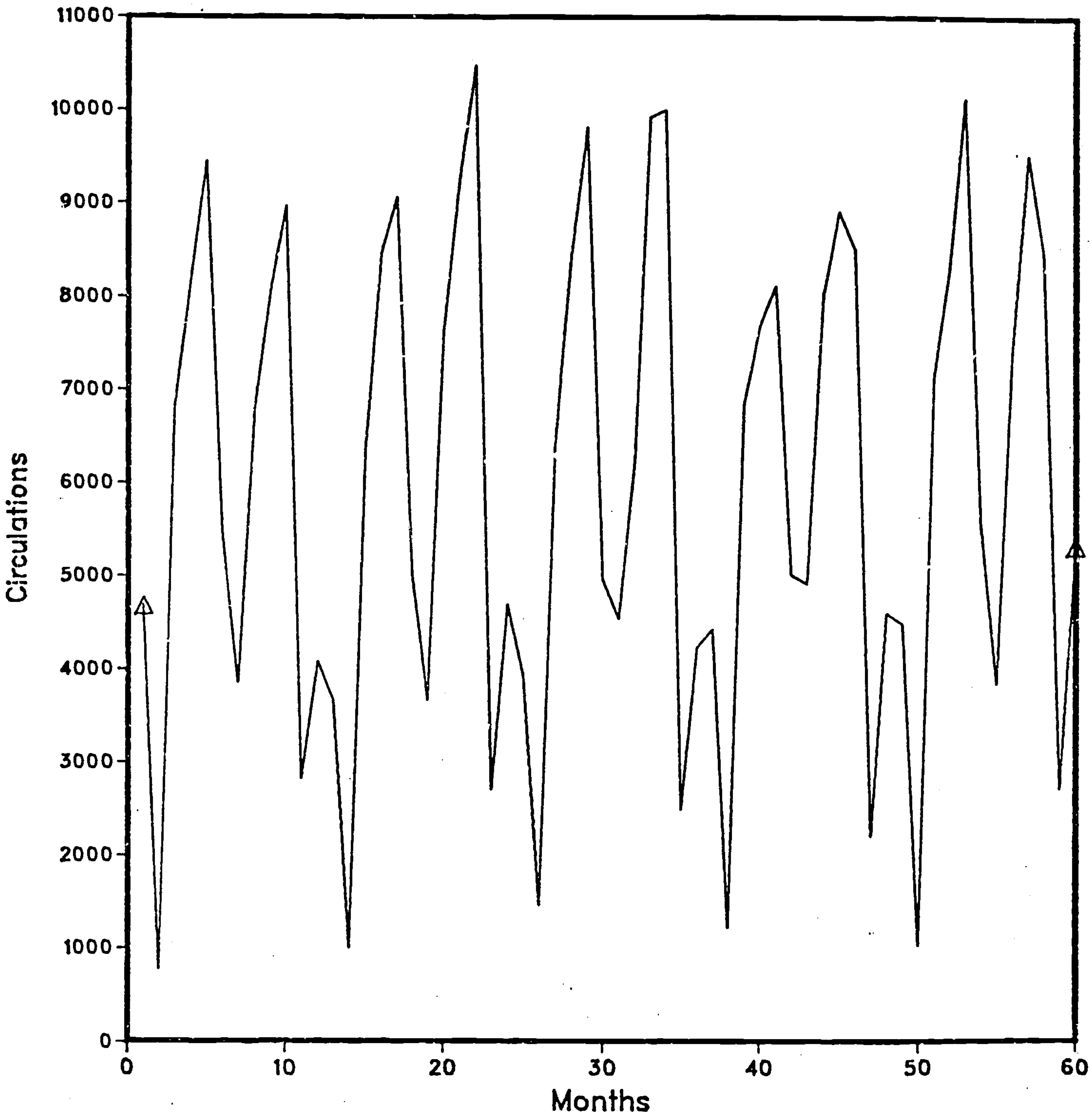
Library 39
Monthly Circulations



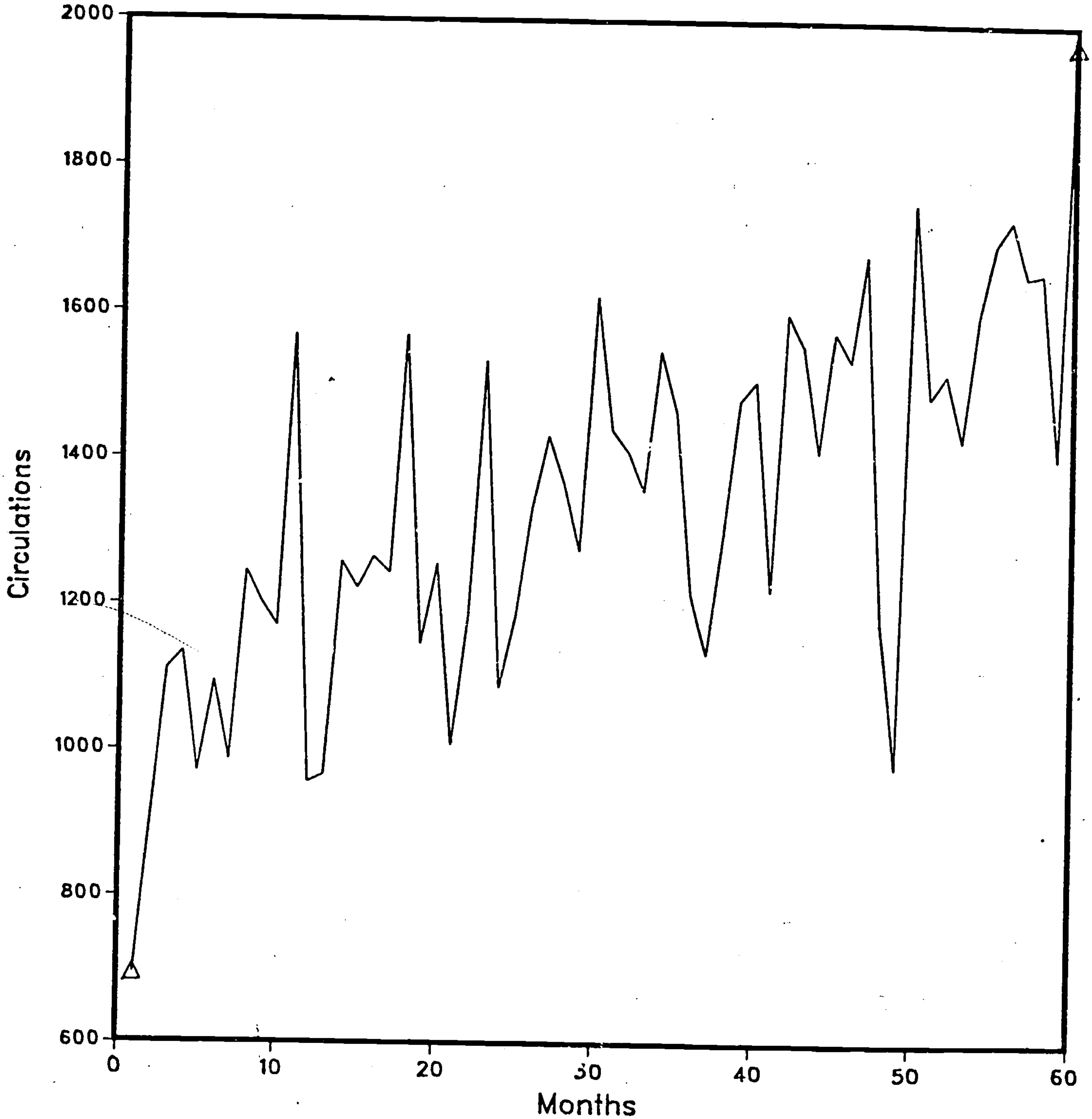
Library 40
Monthly Circulations



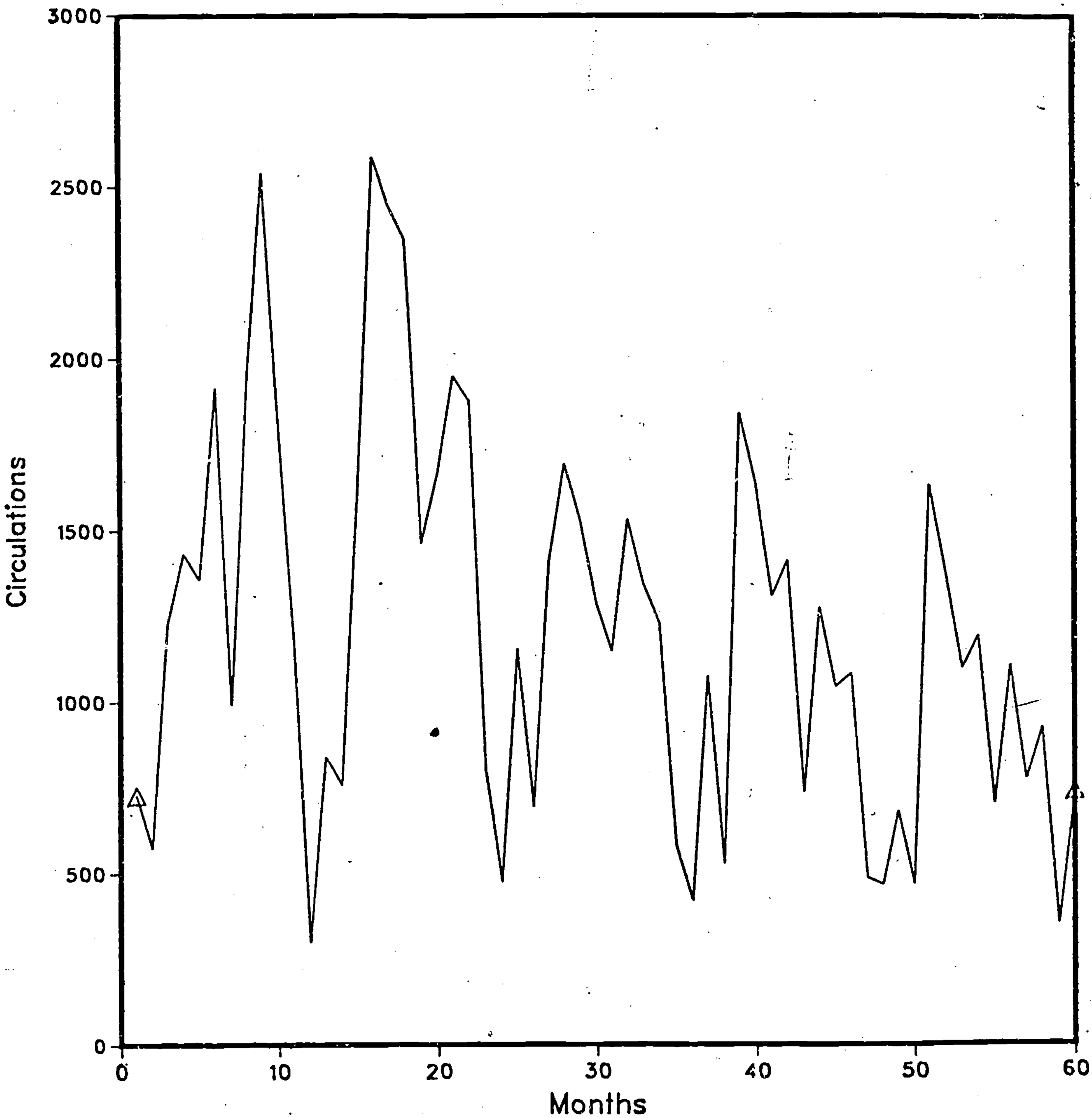
Library 41
Monthly Circulations



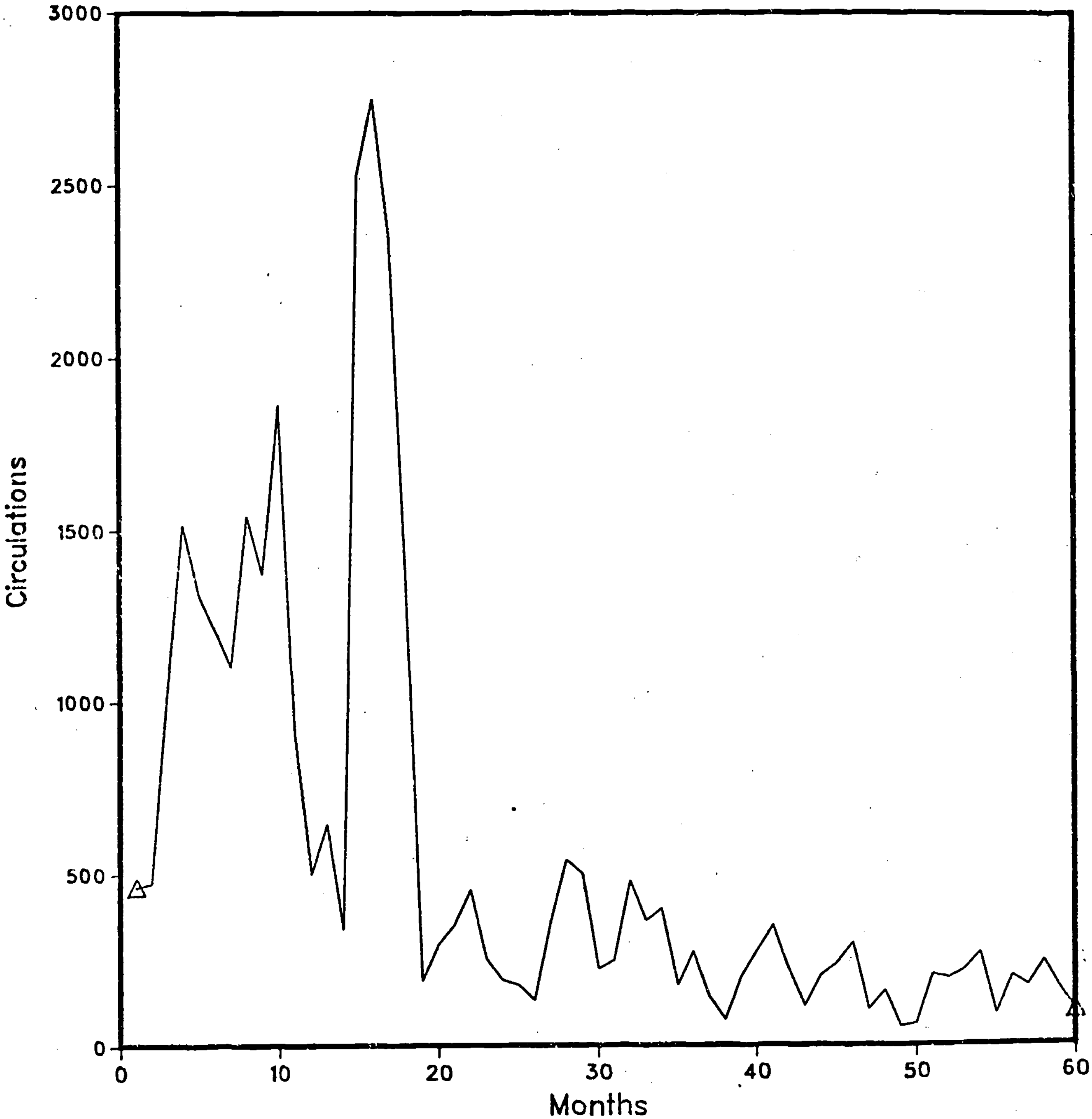
Library 42
Monthly Circulations



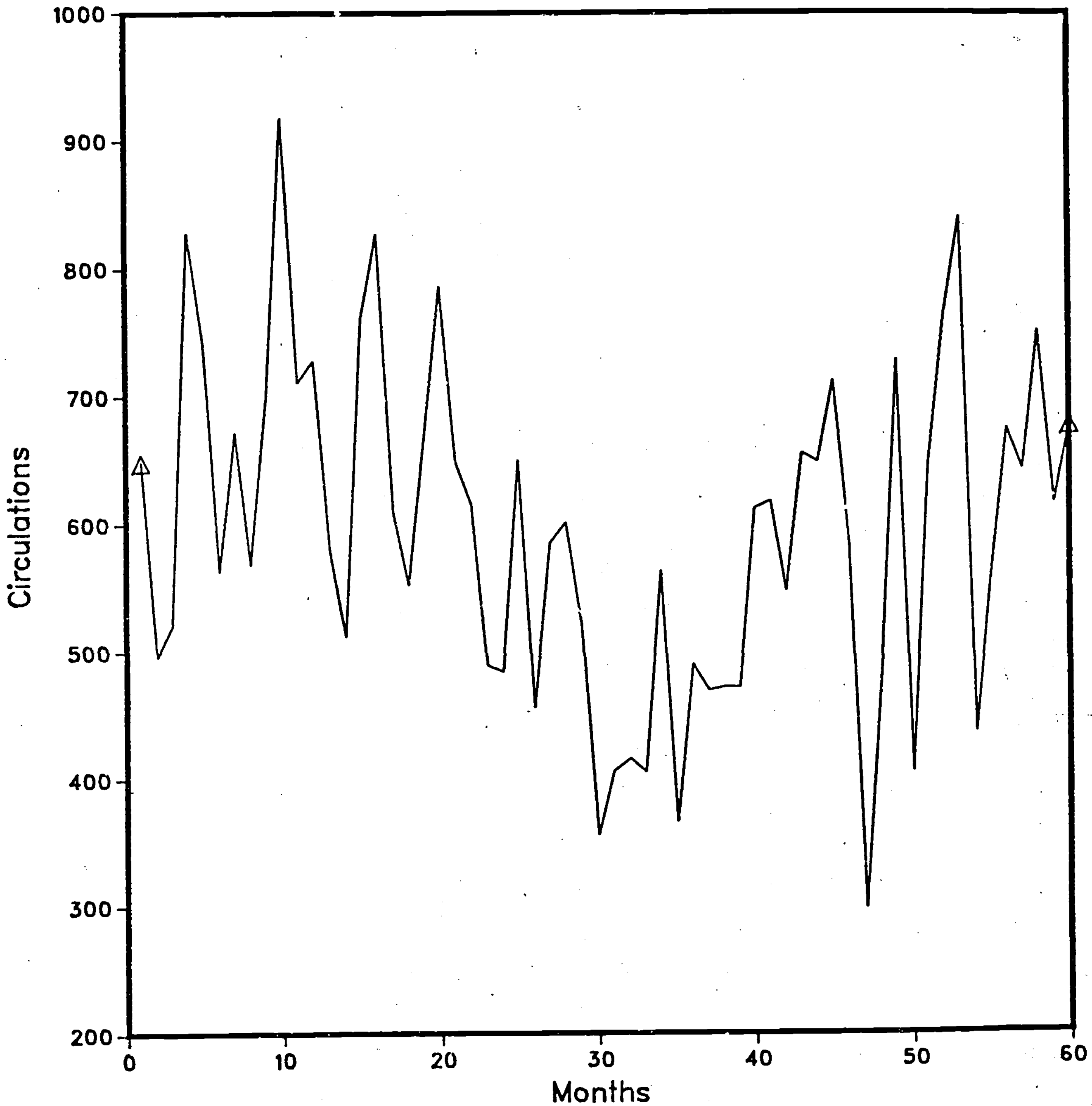
Library 43
Monthly Circulations



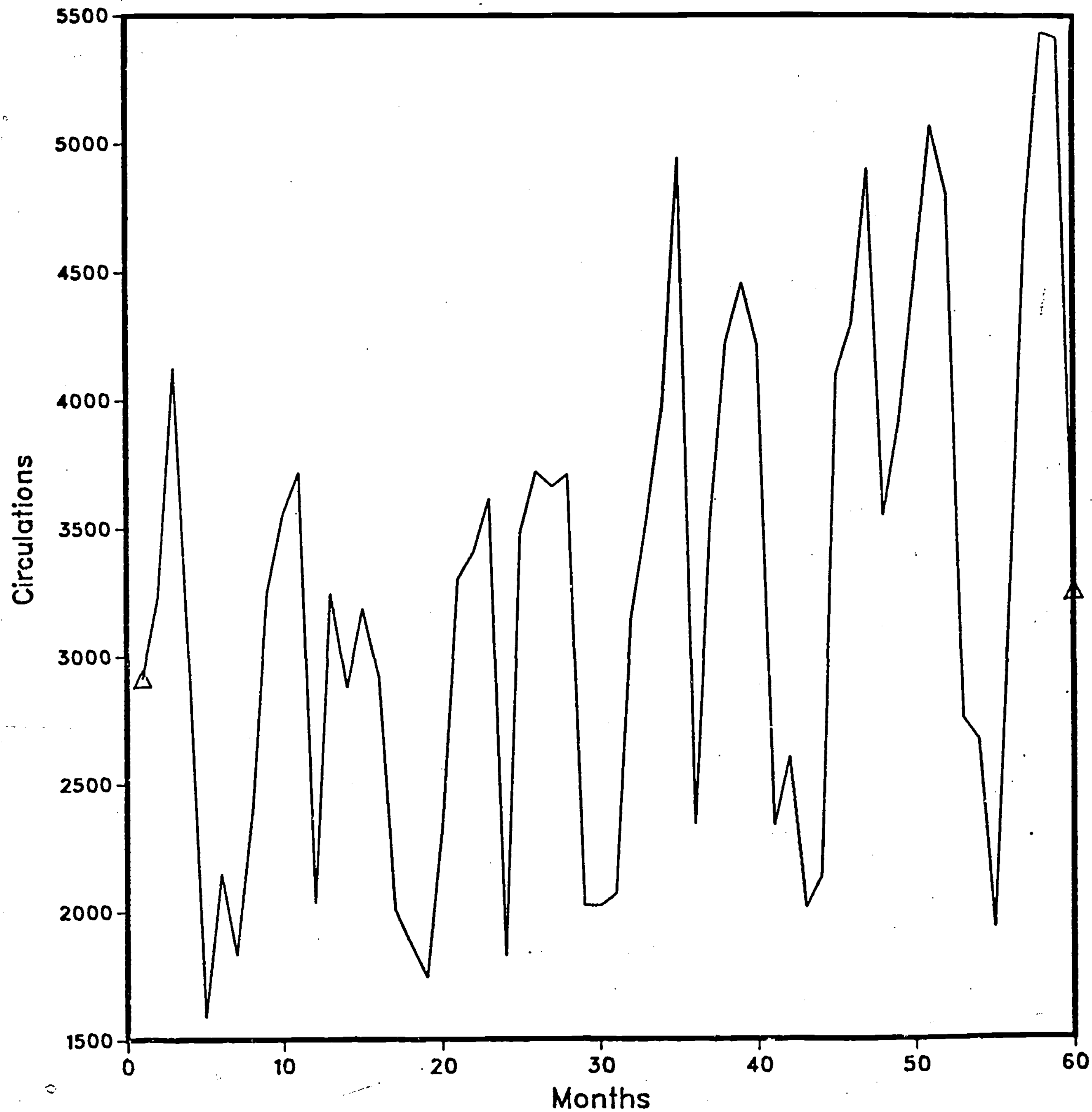
Library 44
Monthly Circulations



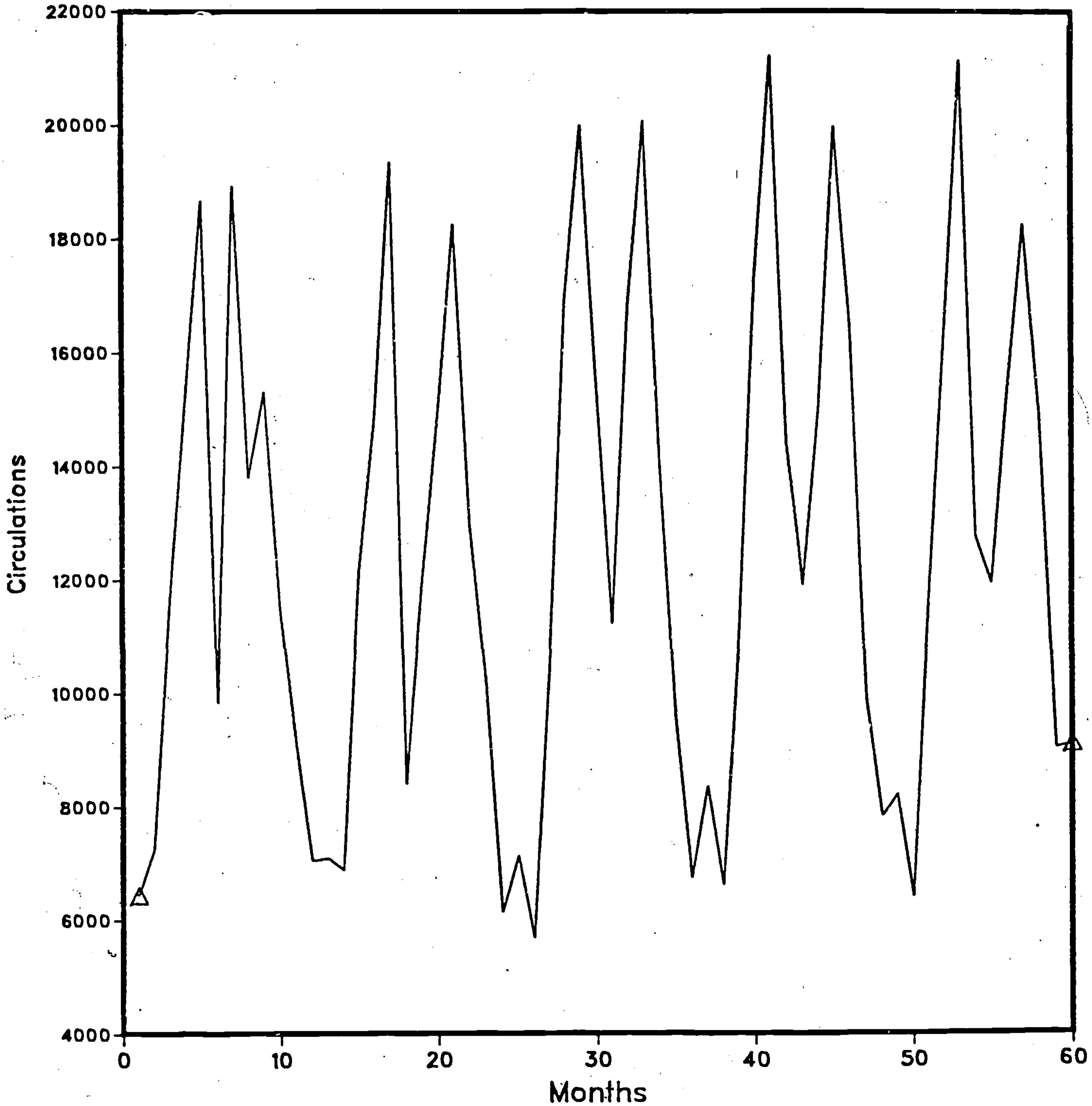
Library 45
Monthly Circulations



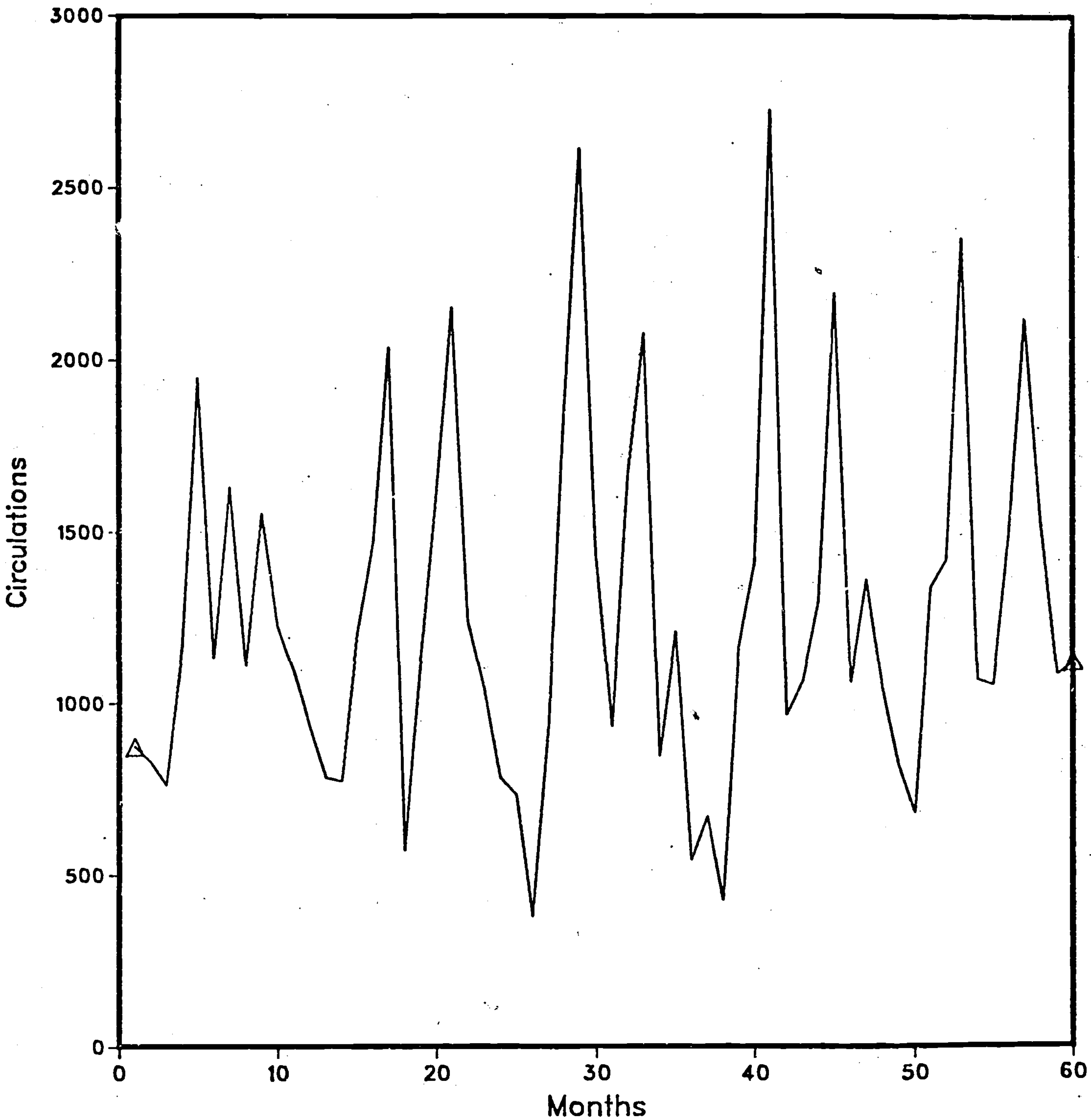
Library 46
Monthly Circulations



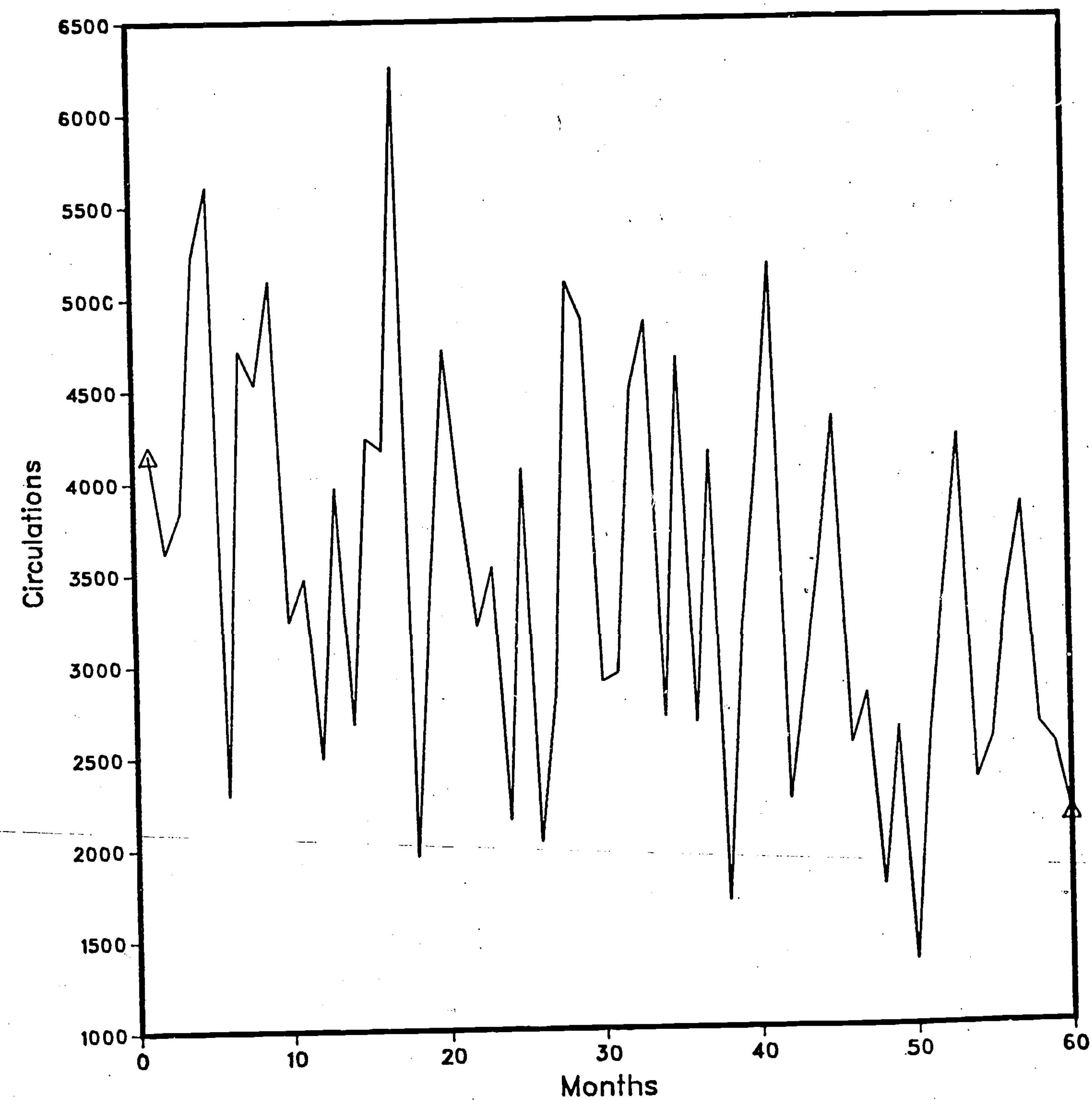
Library 47
Monthly Circulations



Library 48
Monthly Circulations



Library 49
Monthly Circulations



Library 50
Monthly Circulations