

RICHARD W. TRUESWELL

User Circulation Satisfaction vs. Size of Holdings at Three Academic Libraries

In an effort to determine certain facts concerning the relation of circulation satisfaction to collection size, the author sampled the "last circulation date" of stack books and of circulated books in three considerably different kinds of academic libraries. The experience of these three libraries proved to be surprisingly similar. The author speculates concerning potential uses to which such data might be profitably applied.

THE LIBRARY HAS many characteristics similar to the problem and decision criteria associated with the typical industrial inventory problem. The industrial inventory problem is really one of supplying the necessary quantity of the proper goods at the required time and at a minimum cost. Similarly, in the case of the library, it is necessary to provide the correct book when the user needs it, and ideally to have a second book available if desired by other users, all at minimum cost.

In library work, some efforts have been directed at determining which books to keep on the shelf, which books to place in temporary or little used storage, and which books to weed. The problem is further complicated by the need to provide additional copies of certain volumes to satisfy their high request rates.

One approach¹ to this problem em-

¹ Philip M. Morse, "Probabilistic Models for Library Operations," *Association of Research Libraries Minutes*, 63d Meeting (January 1964), 9-19.

Dr. Trueswell is Head of the Department of Industrial Engineering in the University of Massachusetts. A portion of the work herein described was supported by National Science Foundation Grant GN-0435.

ploys frequency studies, where counts are made of the number of circulation transactions for certain sampled books. Attempts were then made to relate these frequency counts to book usage and user needs. One of the difficulties encountered in this approach is that the book card seldom gives a complete history of the transactions for a given book. A frequently used book will, of necessity, require a new book card more frequently than other books, and as each new book card is prepared, the old book card is usually destroyed. Thus relatively short transcription histories are usually available for frequently used books.

Another approach² to finding a parameter for analyzing book usage has involved examining the last or previous circulation date for the book. Books may be categorized by this parameter insofar as there are dates available at which time they previously circulated or, in the case of new books, the time

² Richard W. Trueswell, "Determining the Optimal Number of Volumes for a Library's Core Collection," *Libri*, XVI (1966), 49-60; Richard W. Trueswell, "A Quantitative Measure of User Circulation Requirements and Its Possible Effect on Stack Thinning and Multiple Copy Determination," *American Documentation*, XVI (January 1965), 20-25; Richard W. Trueswell, "Two Characteristics of Circulation—and Their Effect on the Possible Implementation of a Punched Card Circulation Control System," *CRL*, XXV (July 1964), 285-91.

they entered the library. Circulation requirements may then be measured according to this parameter.³

Figure 1 is a cumulative distribution function representing a sampling of circulation and of stack holdings taken at the Deering Library of Northwestern University.⁴ It should be noted that the plot of the circulation sample describes the circulation according to the last circulation date parameter. As noted, the plot reveals the percentage having the last circulation date *not* within the time period indicated; *i.e.*, the percentage of current circulation not previously circulated within time "T." For example, only about 7 per cent of the circulation sample had not previously circulated during the past five years. The stack sample curve represents comparable data taken from a sampling of books in the stacks.

Using both curves in Figure 1, it is possible to make such remarks as the following: at the five-year mark on the plot, approximately 7 per cent of the circulation sample had last circulation dates not within five years, or in effect were taken from a subset or portion of the stack population that comprised approximately 40 per cent of the stack holdings; expressed differently, 93 per cent of the circulation came from 60 per cent of the holdings. As a second example, at the two-year mark the statement could be made that 20 per cent of the stack holdings satisfied approximately 70 per cent of the current circulation sample.

From the information provided in Figure 1 it is possible to make statements about that portion of current circulation that is satisfied by a given per cent of stack holdings. Figure 2 repre-

sents a combining of the two curves found on Figure 1. Thus we have determined points on the curve in Figure 2 referring to their corresponding values on the stack sample curve and the circulation sample curve in Figure 1 for a given time period. For example, in Figure 2 we see that approximately 50 per cent of the holdings satisfy 90 per cent of the circulation requirements.

The data in Figures 1 and 2 represent a very small portion or subset of the holdings at the Deering Library at Northwestern University. The stack sample is approximately 1.7 per cent of the holdings in the 820's.

Additional work has been in progress at the Goodell Library of the University of Massachusetts and also at the library of Mount Holyoke College.⁵ The study at the Goodell Library covered one and one half months and approximately five thousand circulation transactions. These circulation data are plotted on Figure 3 along with stack holdings data and provide some comparison with the very limited data from Northwestern University. Note for simplicity the plots now are last circulation date "within" the time period rather than "not within." Here, for example, 50 per cent of the holdings satisfy approximately 90 per cent of the circulation requirements. Figure 4 shows a curve similar to Figure 2 but applying to the University of Massachusetts. The U/Mass stack sample was approximately $\frac{1}{2}$ of 1 per cent of the library's holdings. The sampling procedure leaves some question as to the randomness of the sample because selection from a fixed number of shelves and a fixed indentation on each shelf was used.

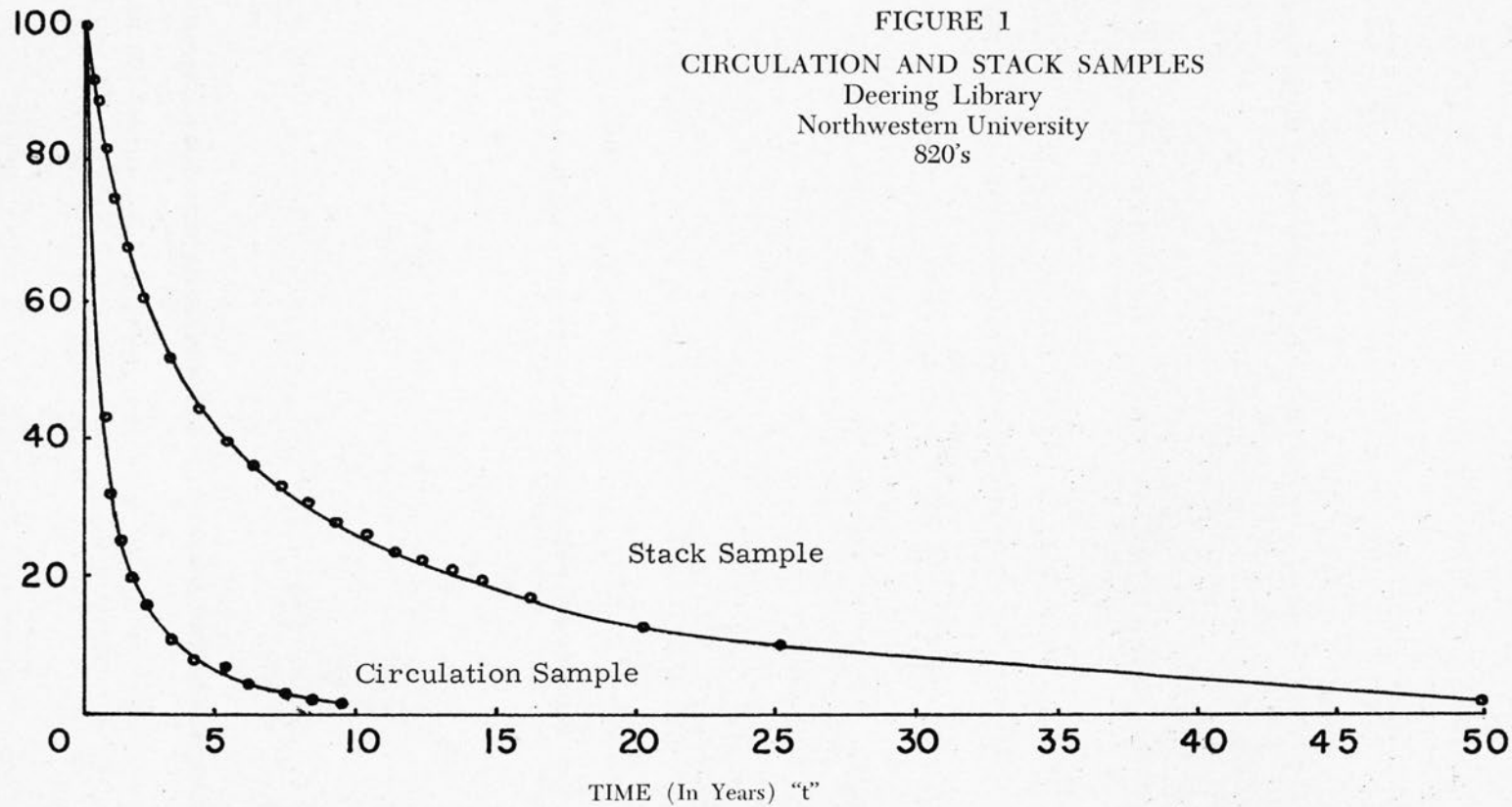
Figures 5 and 6 contain a plot of the last circulation date for the Mount Holyoke College library circulation as recorded over a four-month period. Approximately nine thousand transactions

³ Trueswell, *Libri*, *op. cit.*; Trueswell, *American Documentation*, *op. cit.*; Richard W. Trueswell, "User behavioral patterns and requirements and their effect on the possible applications of data processing and computer techniques in a University Library." (Doctoral dissertation. Evanston, Illinois: Northwestern University, 1964).

⁴ Trueswell, *Libri*, *op. cit.*

⁵ *Ibid.*

% HAVING LAST CIRCULATION DATES NOT WITHIN TIME "t"



% OF CIRCULATION SATISFIED

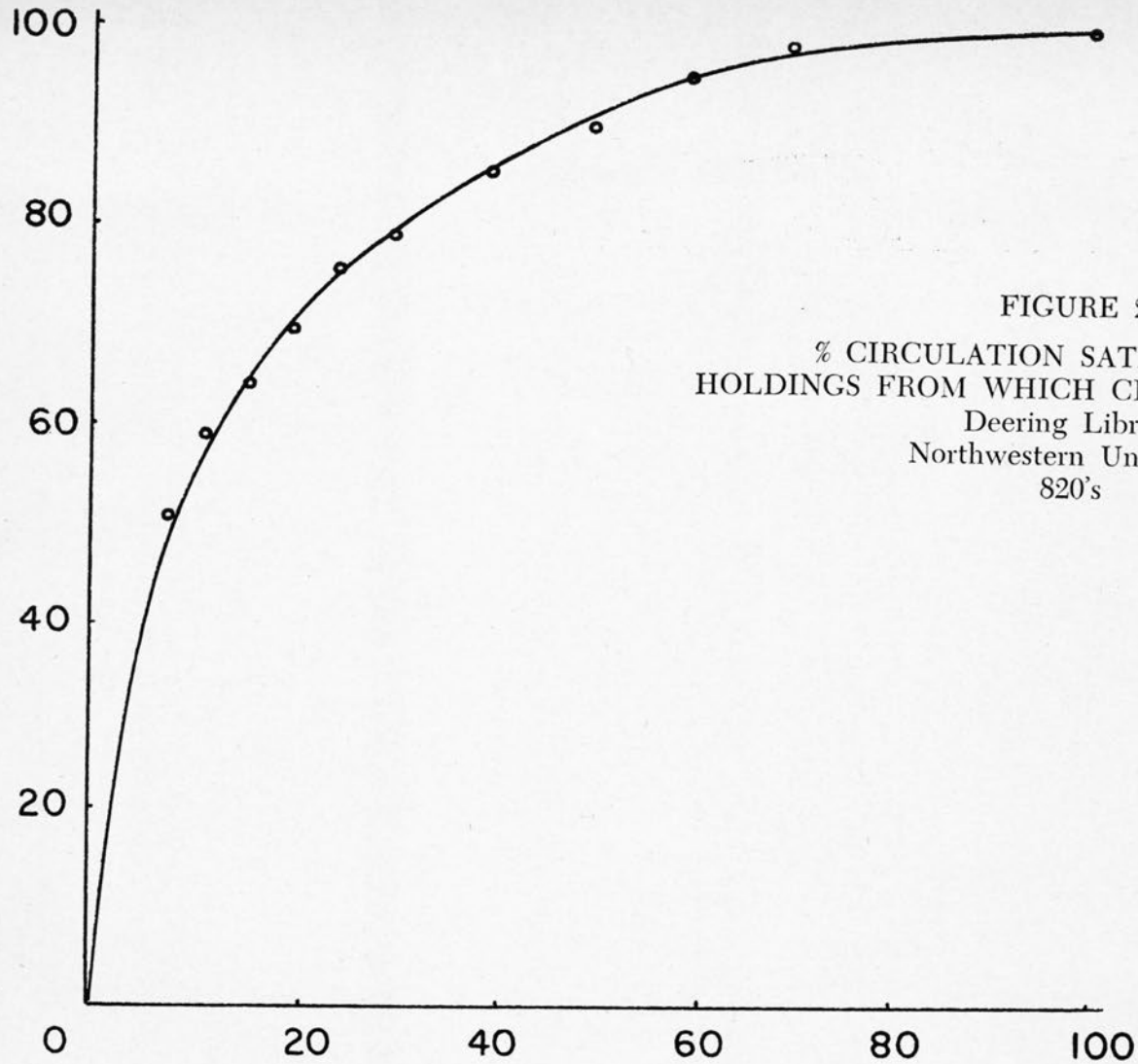


FIGURE 2
% CIRCULATION SATISFIED VS. %
HOLDINGS FROM WHICH CIRCULATION TAKEN
Deering Library
Northwestern University
820's

% OF HOLDINGS SATISFYING CIRCULATION

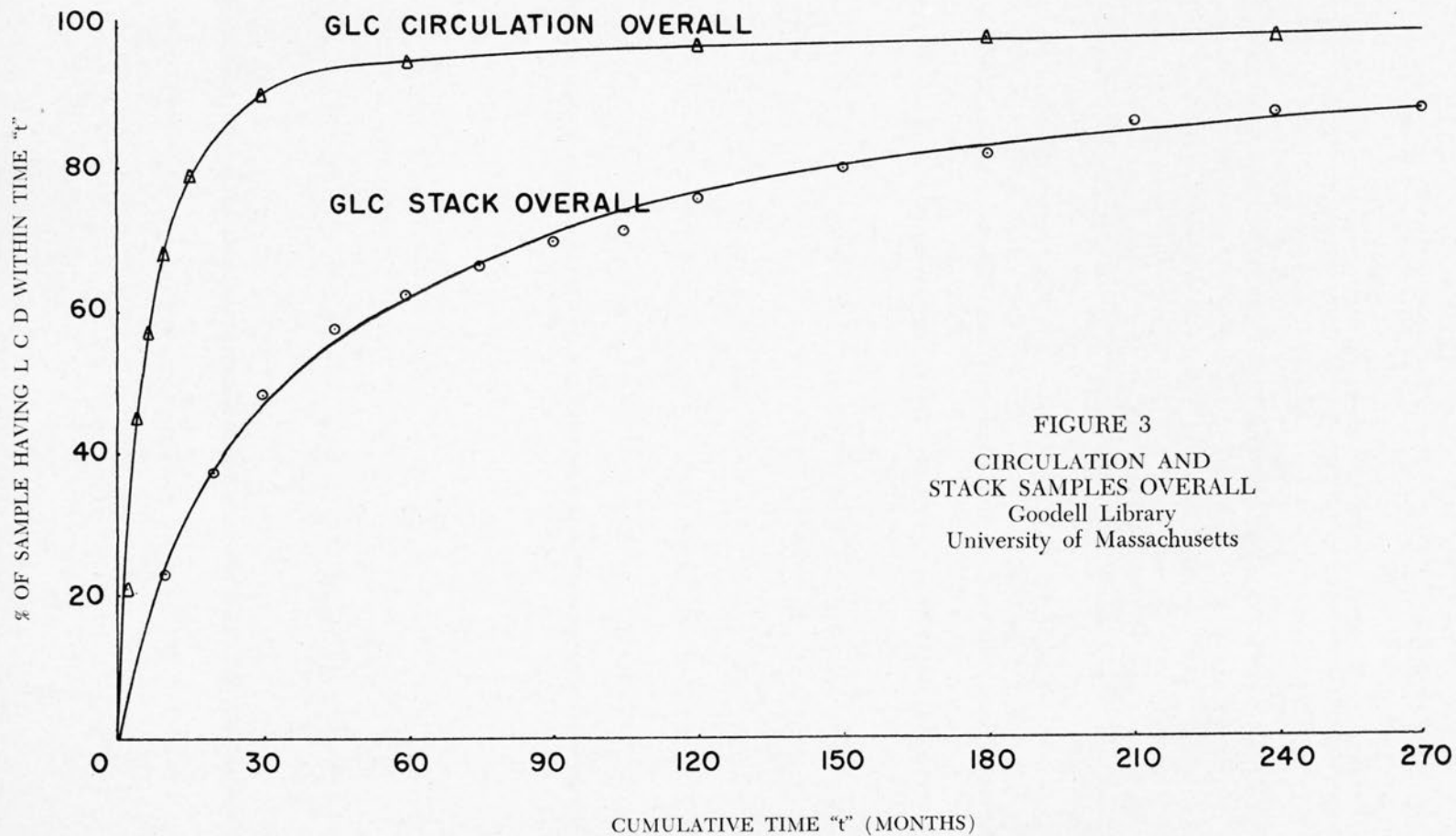


FIGURE 3
CIRCULATION AND
STACK SAMPLES OVERALL
Goodell Library
University of Massachusetts

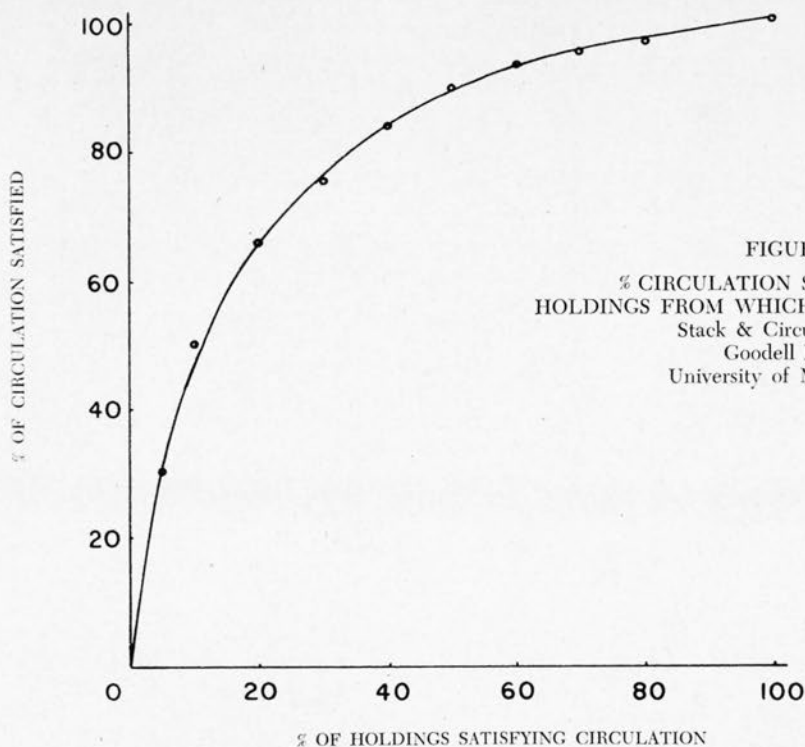


FIGURE 4
 % CIRCULATION SATISFIED VS. %
 HOLDINGS FROM WHICH CIRCULATION TAKEN
 Stack & Circulation Data
 Goodell Library
 University of Massachusetts

were recorded in this sample. Also plotted on Figure 5 for comparison is the last circulation date information for the circulation sample from the Goodell Library and the information from the Deering Library at Northwestern University (over a relatively short period of about one or two weeks).

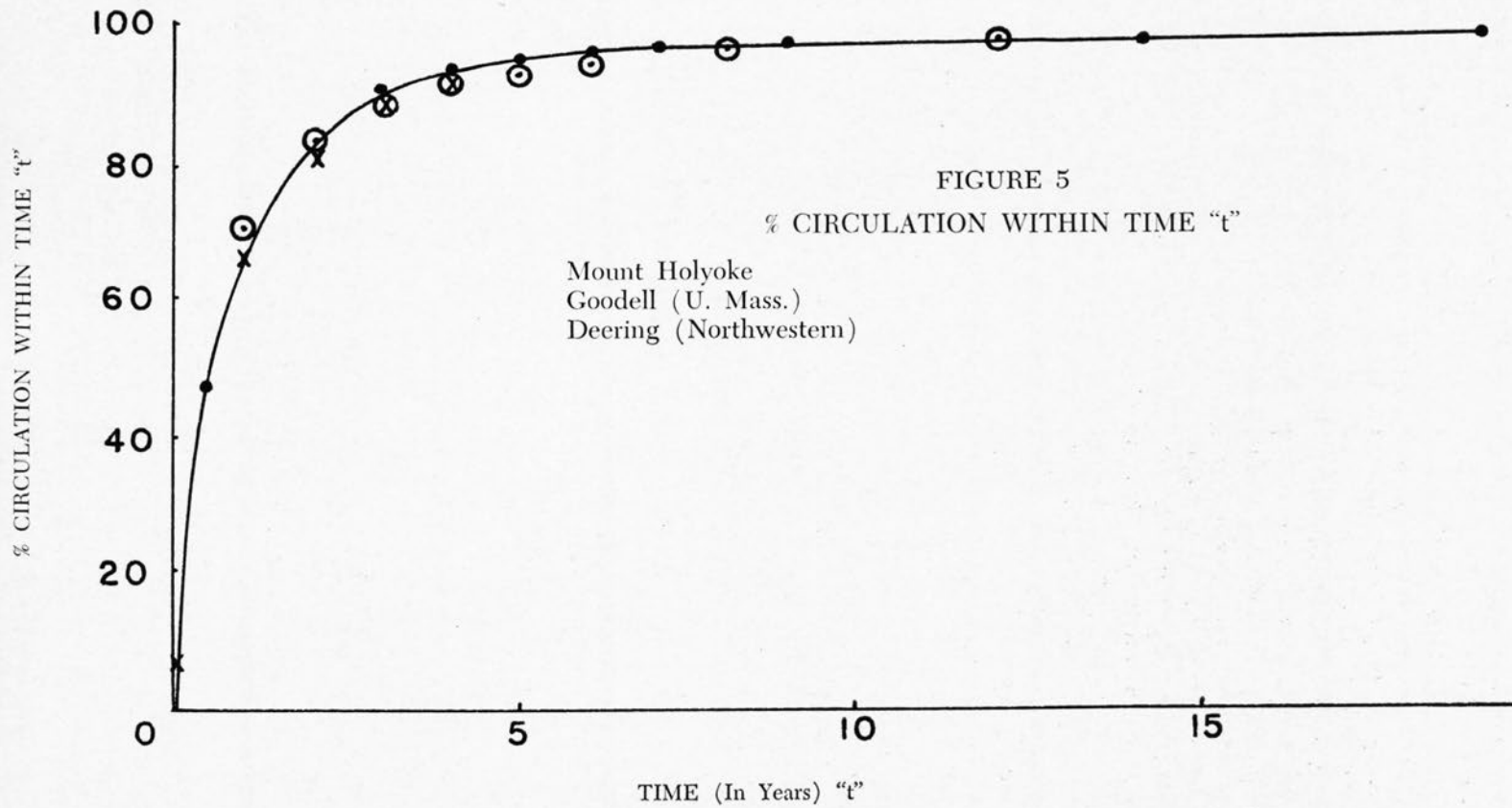
It is interesting to note the similarity of these plots, as they represent the experiences of three different libraries with widely various collections insofar as total holdings are concerned. The shape of the curves indicates that there are a small number of books at each library which circulate very frequently, and suggests that these books could comprise a core collection.

The stack sample at the Mount Holyoke library was also taken during the summer months when circulation was at a minimum. These data (approximately 1 per cent of the holdings) were then plotted as shown in Figure 6. The stack sample data were then plotted against

the circulation sample data to obtain the plot shown in Figure 7, entitled "Percentage circulation satisfied vs. percentage holdings from which circulation taken." This curve is similar to Figures 2 and 4 and illustrates the percentage of the holdings satisfying a given percentage of circulation. For example, we note from Figure 7 that approximately 85 per cent of the circulation is satisfied by 40 per cent of the holdings. It should be noted that these data are quite different from the data collected for the 820's at Northwestern University. The 820's data were taken from a relatively small sample and were limited to a narrow subject category.

It is interesting to observe from Figure 7 and Figure 4 that the plots for the University of Massachusetts library and Mount Holyoke College library are somewhat comparable.

Because of this kind of information, it is now possible to define core collections at any desired percentage of circulation



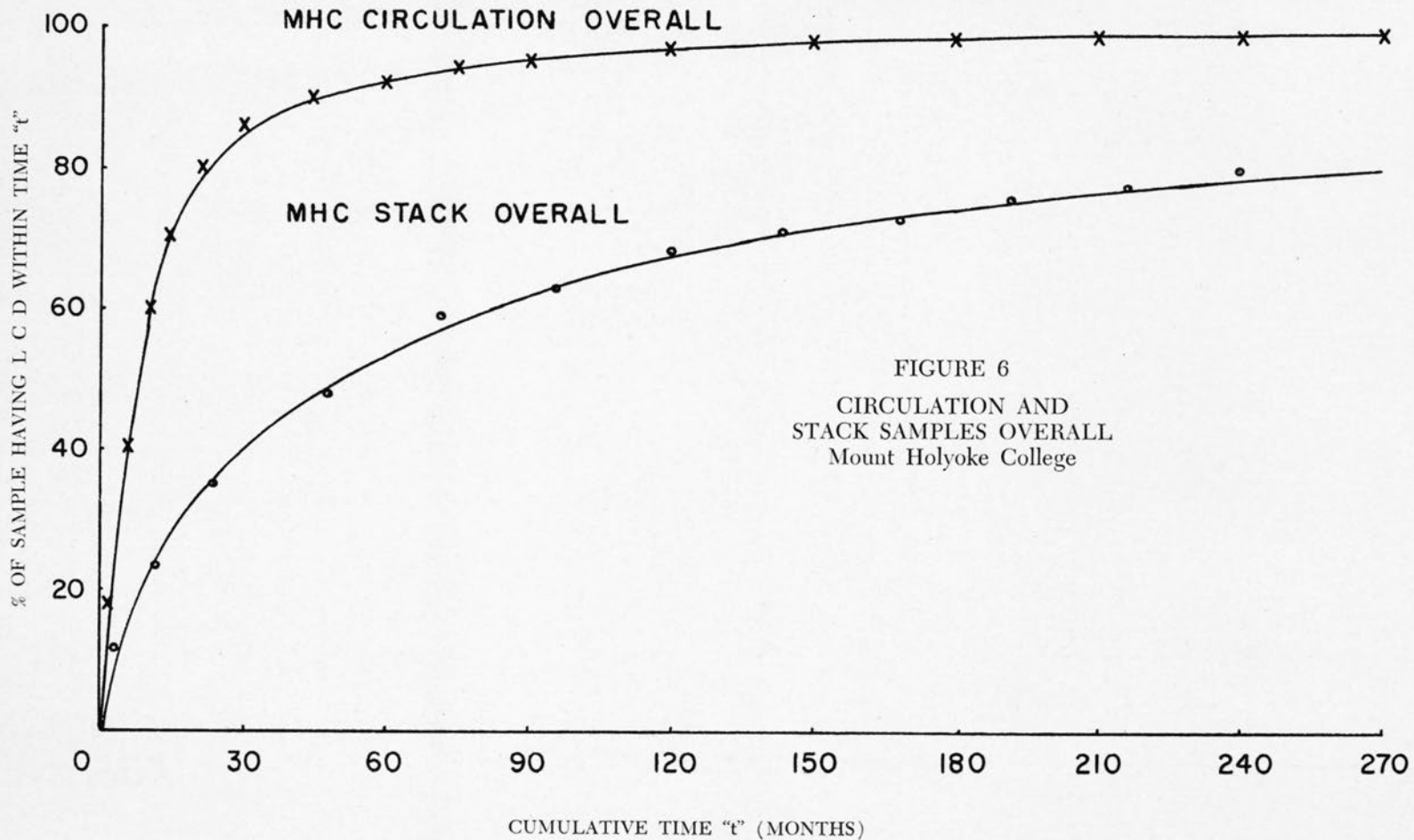


FIGURE 6
CIRCULATION AND
STACK SAMPLES OVERALL
Mount Holyoke College

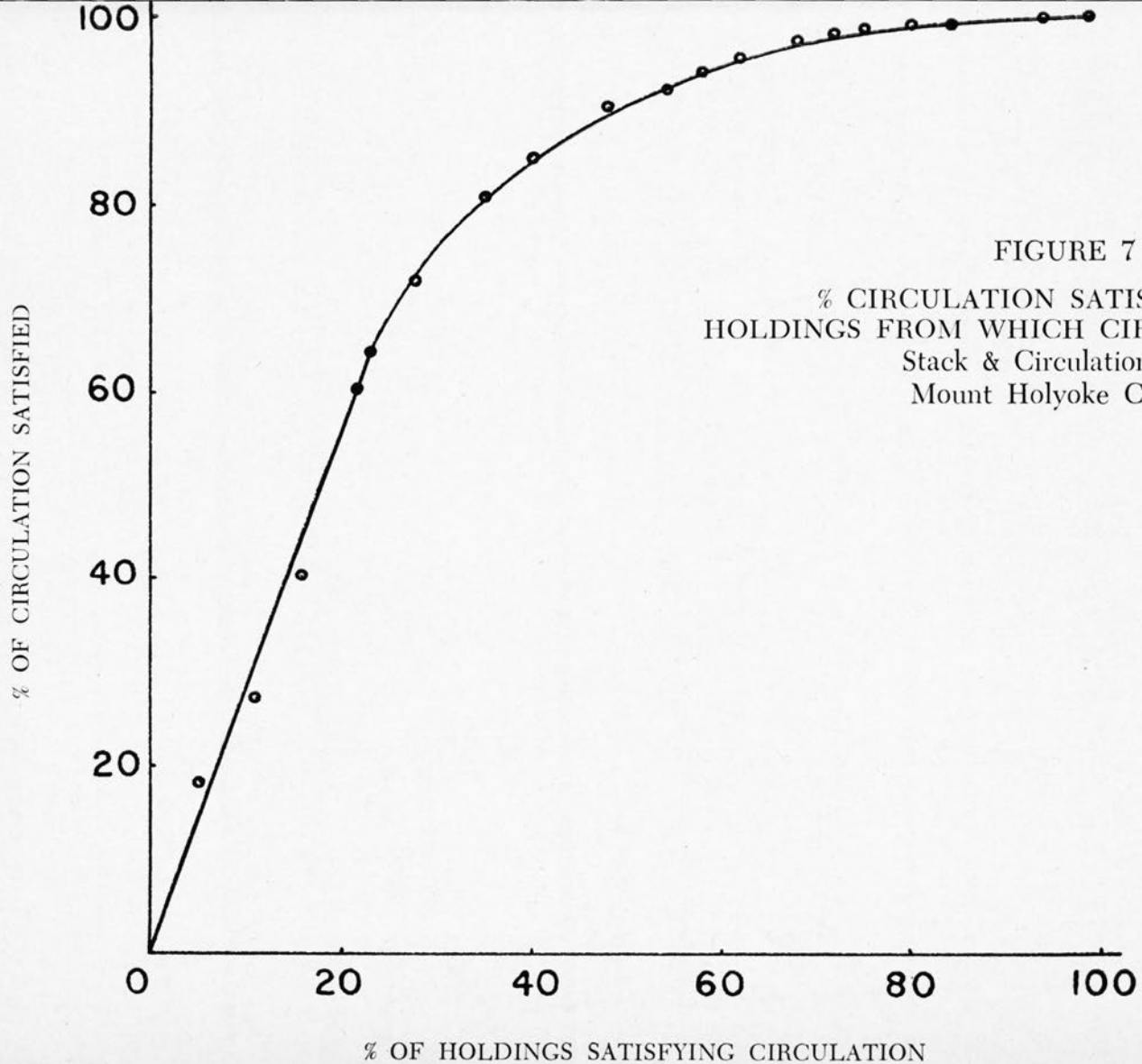


FIGURE 7
% CIRCULATION SATISFIED VS. %
HOLDINGS FROM WHICH CIRCULATION TAKEN
Stack & Circulation Data
Mount Holyoke College

satisfaction. For example, in Figure 7 we could define a core collection comprised of approximately 60 per cent of the current holdings that would satisfy over 94 per cent of the current circulation requirements. Such a core collection could be identified by its last circulation date as taken from Figures 3 and 4.

The author is presently exploring the possibility of predicting the stack holdings distribution (as shown in Figures 3 and 6) utilizing only circulation data. This would require a computer simulation technique whereby the cumulative distribution function would be considered as not changing over time. Circulation rates which do vary over time would be used in this simulation.

If it were possible to predict the stack holdings distribution for a given library using only circulation data from that library, it might then be possible to make statements about the size and circulation satisfaction rates of a core collection. A possible inference might be the conclusion for a given class or category

of library collection, that the size of the core is a function only of the circulation per month or, in effect, only the circulation rate. Thus the size of the core collection for a given library user population would be a function only of the circulation rate expected from that population. The actual percentage of total holdings serving as a core collection (defined at a given satisfaction level) might also be used as an index number to measure the "use" of the library holdings.

More work must be done in evaluating the use of the last circulation date cumulative distribution as a parameter for defining and predicting library circulation patterns. Specifically, this statistic might be used for identifying those books requiring multiple copies; for predicting conversion work loads for changes in classification systems or circulation systems, in studies of high use and low use books, for stack thinning, and as a guide in the selection of titles for new libraries.

■ ■

