RESERVOIR CAPACITY FOR PERIODIC-STOCHASTIC INPUT AND PERIODIC OUTPUT

by

Kedar Nath Mutreja

September 1976

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
Fort Collins, Colorado

86

RESERVOIR CAPACITY FOR PERIODIC-STOCHASTIC INPUT AND PERIODIC OUTPUT

by
Kedar Nath Mutreja

HYDROLOGY PAPERS
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80523

September 1976

No. 86

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The writer wishes to express the gratitude to his advisor and major professor, Dr. V. Yevjevich, Professor of Civil Engineering, for guidance during the writer's doctoral studies. Thanks are extended to Dr. D.C. Boes, Associate Professor of Statistics, for suggestions, and to other members of the writer's graduate committee, namely Dr. D.A. Woolhiser of the Agricultural Research Service, Affiliate Professor of Civil Engineering Department, and Dr. M.M. Siddiqui, Professor of Statistics.

Appreciation is extended to fellow students, especially Norio Tase, Jerson Kelman and Douglas Vargas, for the opportunities of fruitful discussions in many areas of research activities.

Gratefully acknowledged is the support received from Colorado State University in the form of a graduate research assistantship, under the National Science Foundation grant ENG 74-17396, with Dr. V. Yevjevich as principal investigator.

ABSTRACT

A methodology is developed for designing capacities of *large* reservoirs with sufficiently *high* levels of development by using the concept of maximum deficit rather than the range, and short-interval flow records, particularly daily flows.

Daily river flows, composed of periodic and stochastic components, are inputs into reservoirs. The output is assumed to be a deterministic process, either constant or periodic.

The study uses both the analytical and the data generation methods in computing the required storage capacities for regulating the periodic-stochastic inputs to deterministic outputs. The approach used in determining the mean required storage capacity, named in this study the total storage, is to divide the total storage, defined as the expected maximum deficit of the net input to a reservoir, into the difference storage and the stochastic storage, with: total storage = difference storage + stochastic storage.

Difference storage is due to periodic components of both input and output. The advantage of its use is that it is approximately constant for large sample sizes. It can, therefore, be estimated for different sample sizes by generating a relatively short series of daily flows.

Stochastic storage, defined as the expected maximum deficit of the stochastic net input, is estimated analytically by using a coefficient, estimated by generating short series of daily flows.

It was found that the parameters of asymptotic distributions of statistics of partial sums can be obtained by equating the higher-order autoregressive models to an equivalent first-order autoregressive model with its ρ_1 equal to the sum of all the autoregressive coefficients of the higher-order model. This is valid for all storage problems of practical interest.

The methodology developed for determining the total storage by its decomposition into the difference storage and stochastic storage has been applied to a simulated problem of determining the storage capacity of a reservoir with the economic life of 100 years to be constructed at a site for which there are 40 years of daily flow data (the Oconto River, near Gillett, Wisconsin).

FOREWORD

Planning, decision making on the size, and management of water storage capacities represent research topics which have been studied in the past from two points of view, practical and theoretical. About a century of modern planning and construction of water storage capacities in different forms in various countries has produced a practice, based on experiance, on how to plan, determine the size and operate these water storage capacities. In the practical, engineering approach, basically the historic water input data are used in such a way that the storage capacity is determined assuming the operation of the reservoir has started at the beginning of data collection and lasted until the end of available series. The storage problems are studied with the premise that a given water draft regime, satisfying a given water demand or solving complex water resources problems by regulation, is accomplished by that capacity during the historic data period. On the other side, theoretical approaches have been developed for studying various storage problems, and not only water storage but also for storing goods, oil, gas, and other materials, for which either the input, or the output, or both are governed by the law of chance. The theory of storage has become a part of mathematics, or especially a part of probability theory, stochastic processes and mathematical statistics. The storage problems have received several theoretical treatments in the past, especially in the last 30-40 years. In recent times, the theoretical approach to water storage problems has produced many analytical solutions by several schools of probabilists and statisticians.

It may not be exaggerated to state that the above two approaches, one based on experience and practice and the other on theoretical analysis, have not been yet integrated. No real bridge has been made between the results of these two approaches. The practice of water storage planning and operation has been related for too long a time only to historic data series. Since the future samples may show the need for different storage capacities than those obtained from the historic samples, there is some concern among specialists at present about the correctness of decisions made on the size of large storage capacities in comparison with the mean annual inflow when based only on this empirical approach. On the other hand, to tackle mathematically the complex inputs, outputs, and the changes in boundary and initial conditions of water storage capacities, it is necessary to make many simplifications that are mathematically imposed in order to produce analytical solutions.

The most difficult problems of the theoretical analysis of water storage capacities have resulted from periodicities in parameters of hydrologic input and water demand time series. In the most general case, not only the mean and standard deviation are periodic parameters, but also the other parameters may be proven to be periodic, such parameters as coefficients of dependence models of stochastic components and higher-order moments as the skewness and kurtosis coefficients. Even in the simplest case for which only the mean and the standard deviation of inputs and outputs are periodic, the theoretical approach becomes relatively difficult to implement. When the realistic inputs and outputs of storage capacities are used, it is obvious that the bridge between theory and practice can not be made easily. This does not result only from the periodic-stochastic character of inputs and outputs but also because of usual trends in water demand and the non-stationarity of available storage capacity due to sedimentation of reservoirs with time. Efforts undertaken by many research groups around the world to make a bridge between theory and practice, and to use theory to make better planning, decision making on the size, and operation of water storage capacities, did not yet produce the satisfactory and generally acceptable results.

Several studies in Hydrology and Water Resources Graduate and Research Program of Civil Engineering Department at Colorado State University, especially studies in the form of Ph.D. dissertations, have been undertaken with the basic objective in mind to make contributions of research results for a better bridge between theory and practice. The Ph.D. dissertation by Kedar Mutreja, which is the subject of this hydrology paper, should be looked at from the point of view of making a contribution toward that bridge.

The problem of periodicity in the mean and the standard deviation is approached in the paper in such a way as to enable the division of the expected or total, finite storage capacity into a deterministic and a stochastic part. This is similar to the decomposition of hydrologic time series into periodic parameters as deterministic functions of time and a stationary stochastic component. The expected storage capacity or the total storage is divided in this paper into the stochastic storage which is the result of the random variable $(\sigma_\tau \varepsilon)$ as the product of periodic standard deviation (σ_τ) and the stationary stochastic component (ε), and the difference storage, as the difference between the total storage and the stochastic storage. Therefore, the effects of periodicity in parameters are divided and studied basically as the effects of periodicity in the means of inputs and outputs by using the difference storage, and the effects of periodicity in standard deviation and the stochastic component by using the stochastic storage.

This paper presents the investigation of effects of various parameters on the difference storage. These parameters are: periodicity in the input mean, general standard deviation of inputs, first serial correlation coefficient of the stationary stochastic component, periodicity in the output mean, and finally the sample size n. The difference storage converges to a constant with an increase of the sample size. The property of difference storage is conceived as the deterministic part in the total storage capacity. The stochastic storage is studied as a function of periodic standard deviation, linear dependence of stationary stochastic component and sample size.

A distinction is made in the paper between the full and the partial utilization of regulated river flows. The full utilization means the 100% use of the available water to be regulated. The partial development represents the use only of a fraction in the regulated form of the total average flow, with the difference being the unused, spillover water. It has been shown by many researchers in the past that the necessary storage capacity decreases with a decrease of the degree of development in water utilization. The theoretical studies of water storage have been mainly carried out for the full development, namely for the output mean being a constant equal to the input mean, or for the outflows varying with time but having the mean equal to the input mean. The theoretical studies of partial development, which temporarily sacrifices a percentage of the total available water but is a realistic case in practice, have received a relatively limited coverage in the past. Therefore, if the objective of flow regulation is a partial development of the water resources potential, the total storage capacity must be a function of the level of development. By using the method of generating new samples, the difference storage for large sample sizes in days oscillates around a constant, with fluctuations only due to sampling variation, and to an eventual small effect of periodicity in inputs and outputs.

Stochastic storage, or for that matter the total storage, is investigated by using the concept of the maximum deficit (or the maximum depletion). If it is assumed that the finite storage capacity of a reservoir is full at some point in time, the water spills over as long as inflows are higher than outflows. The peak at the cumulative curve of these differences is reached when the output is equal to the input and the output starts to be greater than the input. Then the depletion of the reservoir storage begins. The maximum depletion or deficit as the difference between this peak and the next lowest point at the cumulative curve of differences between input and output, even with the new ascending branch of this cumulative curve exceeding the previous maximum, represents the largest deficit or the largest depletion as the necessary storage capacity. The maximum differences between the successive maxima and minima represent a random variable, important for decision making on the size of water storage capacities. The finally selected capacity can be determined from the probability distribution of that variable by an optimization analysis. The risk is always involved that the selected storage capacity will not be able to supply the depletion water volume in all the samples of a given size. In general, the maximum deficit is a variable smaller or equal to the range. The range is defined as the difference between the maximum and the minimum of the cumulative sum of input minus output for a given sample size. The mean deficit is then always smaller or equal to the mean range.

Because the range and the deficit are defined on the cumulative sum of differences of inputs and outputs for a given sample size n, it can be shown that the investigation of asymptotic deficit, range and the other water storage parameters of stationary stochastic components which follow the linear dependence models, can be reduced to investigations of the time independent stochastic component (TISC) inside this stationary process. The dependence is taken into account by a parameter (β) which is the function of all the coefficients of the linear dependence model. For purposes of generating new samples, an m-th order linear autoregressive model can be reduced to the study of the first-order autoregressive model only. In this latter case, the serial correlation coefficient ρ of this first-order model is equal to the sum of all the coefficients in the m-th order autoregressive model. In general, it can be shown that in the study concerning the asymptotic results of the deficit, the range or any other water storage parameter, or the study of distributions of deficit and range of linear dependence models, can be carried out only by investigating the properties of these parameters for the time independent stochastic component (TISC). For every linear dependence model a parameter (β) could be found which helps to relate the necessary storage capacity (deficit or range) of the autoregressive, stationary stochastic components to the storage capacity of their time independent stochastic components.

V. Yevjevich
Professor-in-Charge of
Hydrology and Water Resources Program

Chapter 1
INTRODUCTION

1-1 General

Since nature does not meet the demand for water in time and space, reservoirs have been most important for regulation of surface runoff so as to balance supply and demand. Although man has an experience with reservoirs for the last three to four millennia, and although he has developed economical methods of dam construction for wide variety of sites to create reservoirs, the methods of estimating the proper sizes of dams and reservoirs for the target pattern of demand have been mainly based on "rules of thumb" and "engineering judgement."

The empirical method of analysing the stochastic problems has dominated the engineering practice ever since Rippl applied this approach in 1883, (Rippl, 1883). Rippl's method, based on the mass diagram, assumes that both inflow and outflow are known functions of time. It gives the minimum storage capacity required that no water shortage would occur during the period under consideration. The reliability of results so obtained is limited, because the analysis is necessarily based on a single historical sequence of hydrologic records. The probability is zero that an identical flow sequence would occur again during the active life of a reservoir. Moreover, the length of the historical record is apt to be quite different from the economic life of a reservoir, which in turn is determined not only by social and economic considerations but also by pure physical considerations. Since the required storage capacity obtained by using the Rippl mass-curve method increases with the increase of the length of record, the estimated capacity usually will be incompatible with a design based on the economic project life. Because of only one storage capacity value as the result of mass-curve method, this method does not provide information to a designer in finding out the risk to be taken with regard to water shortages during periods of low streamflows.

The fact of the matter is that one cannot speak of the needed storage capacity of a reservoir in a deterministic sense, because of the stochastic nature of both the streamflow and water demand. The needed capacity for a given sample size is a random variable, necessitating, therefore, the consideration of its distribution, with the expected value and variance of the distribution of this variable being important parameters in the final selection of the storage capacity of a reservoir.

The theory of stochastic processes, applied to design and operation of reservoirs, has recently become one of the most important topics of statistical hydrology. Unfortunately, the storage problems are extremely complex. The complexity depends on the type of required or proposed regulation. For example, if the regulation is of the over-the-year type, the analysis is based on annual streamflows and a given degree of river development or water draft. In dealing with annual flows the assumption of their independence may be sufficiently accurate in many cases. In general, the serial correlation is such that the Markov or linear autoregressive models are needed to describe this dependence (Yevjevich, 1964; Fiering, 1967). The natural annual flows may be considered as stationary stochastic processes. Hence, properties of the random variable of storage capacity may be obtained either by the exact or approximate solutions.

When within-the-year water fluctuation is to be considered in the design of a reservoir capacity, then one has to analyze the problem with either monthly, weekly or daily runoff series along with the respective monthly, weekly or daily water demand series. The analysis requires a consideration of nonstationary stochastic processes, because both time series show periodicity in the mean, standard deviation and often in autocorrelation coefficients, besides the time dependence structure of stationary stochastic components, (Thomas and Fiering, 1962; Roesner and Yevjevich, 1966; Yevjevich, 1971). Besides this the trend variability in water demand and the competition between water users add to the complexity of the whole analysis.

The complexity of the problem is well demonstrated by the fact that in search of a solution to this problem, the engineers like Rippl, Hazen, Sudler and Hurst have, respectively, introduced the concept of mass curve, invented such a useful tool as the probability paper, pioneered methods of simulation, and to raise questions which still remain unresolved.

1-2 Approaches to Investigation of Storage Problems

Approaches commonly used in the design of storage capacities may be classified into three groups: empirical, experimental, and analytical. The empirical approach consists in the application of Rippl's mass curve as described in Section 1-1.

The experimental approach is simply the application of the Monte-Carlo or data generation method. The central idea of mathematical synthesis is to create the periods of high and low runoff that mostly are not present in short historical records. However, these periods, from the view point of probability theory, could be expected to occur in an actual record of sufficient length. The Rippl's method, or a modification of it, is then applied to each flow sequence. The probability distribution of storage capacities is approached by using the relative frequency distribution as its estimate. The more samples are generated, the better is this estimate.

In order to deal with the analytical method a few definitions are needed. Let $\{Z_i\}$ be a sequence of random variables such that $E(Z_i) = 0$ and let

$$S_i = z_1 + z_2 + \dots + z_i;\ i = 1, 2, \dots, n \quad (1\text{-}1)$$

$$M_n = \max\ (0, S_1, S_2, \dots\dots, S_n) \quad (1\text{-}2)$$

$$m_n = \min\ (0, S_1, S_2, \dots\dots, S_n) \quad (1\text{-}3)$$

$$R_n = M_n - m_n \quad (1\text{-}4)$$

with the random variable S_i = the cumulative or partial sum, M_n = the maximum partial sum or surplus, m_n = the minimum partial sum or deficit, and R_n = the range of the partial sums (Fig. 1-1).

Another type of these three statistics occurs when each component of the partial sum is corrected

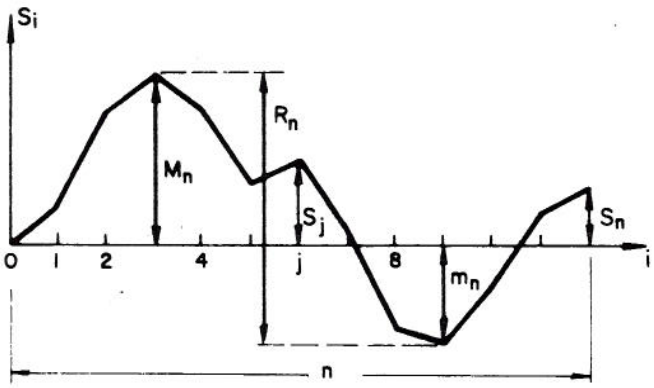

Fig. 1-1. Definition of the Maximum Partial Sum (M_n), the Minimum Partial Sum (m_n), and the Range (R_n).

for the sample mean, $\bar{z}_n$. Therefore, the above random variables will then become

$$S_i^* = S_i - (\frac{i}{n}) S_n \tag{1-5}$$

$$M_n^* = \max (0, S_1^*, S_2^*, \ldots\ldots, S_n^*) \tag{1-6}$$

$$m_n^* = \min (0, S_1^*, S_2^*, \ldots\ldots, S_n^*) \tag{1-7}$$

$$R_n^* = M_n^* - m_n^* \tag{1-8}$$

where S_i^* = the adjusted partial sum, M_n^* = the adjusted surplus, m_n^* = the adjusted deficit, and R_n^* = the adjusted range. These statistics are shown in Fig. 1-2.

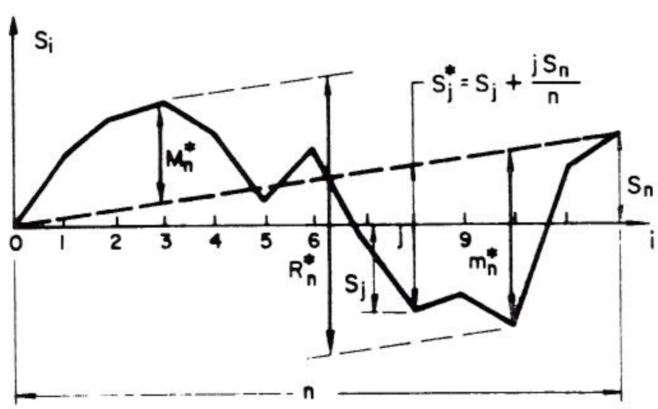

Fig. 1-2. Definition of the Adjusted Partial Sum (S_j^*), the Adjusted Maximum Partial Sum (M_n^*), the Adjusted Minimum Partial Sum (m_n^*) and the Adjusted Range (R_n^*).

Some engineers interpret range as the required storage capacity to avoid both overflows and emptiness of the reservoir. However, this may be valid *only* in the case of *full regulation* or full development. Full regulation of river discharges is equivalent to assuming the expectation of random variable Z_i of Eq. (1-1) equal to zero. But when this expectation is positive then it is a case of *partial regulation* or partial development. Another statistic known as the maximum deficit, with its definition based on the cumulative departures from an arbitrary base value less than or equal to the mean, is used to work out the required storage capacity of the reservoir.

Figure 1-3 gives the plot of cumulative sum of departures for different base outflow values. The summation curve is always studied in conjunction with the inclined axis representing the base value. It is obvious that as the inclination of axis OA changes, different points on the summation curve may become maxima or minima. For instance the vertical distance between points e and f gives the range and between f and g the maximum deficit with respect to the base value OA respectively. It may be noted that the deficits $\{d_i\}$ in Fig. 1-3 are measured from the subsequent peaks higher than the previous peak. However, the maximum deficit for the base values OA_1 and OA_2 are respectively the vertical distances of a-c and b-c. Thus the range as the criterion for storage capacity is obviously not correct at least not for the partial regulation.

The analytical method consists of finding by exact, asymptotic or approximate derivations of various descriptors related to storage capacity design, such as mean, variance and other descriptors of surplus, deficit, and range. Exact general expressions for some of these descriptors are derived only for the case of independent, identically distributed random variables and for the stationary first-order Markov linear model. Similar properties are not available when random variables are nonstationary.

Empirical data generation and analytical methods in solving storage problems were used in the analysis of reservoir storage design and operation in the past (Yevjevich, 1965). However, the advent of computers enabled the generation method to be very attractive. Mathematical methods using the probability theory, mathematical statistics and stochastic processes were tried by many investigators by solving the water storage differential equations under various conditions in the last three decades.

1-3 Description of Model

The solutions of storage problems in case of within-the-year water fluctuation are topics of this study. The need to deal with nonstationary series of inputs and outputs makes the general mathematical treatment of storage problems extremely complex. Hence in practice one falls back on the generation procedure in order to tackle many problems. This study is conceived to use both methods, mathematical and data generation, in computing the necessary storage capacities of complex periodic-stochastic input and output processes.

The study is concerned with the design of storage capacity of an independent reservoir. An independent reservoir is defined as a reservoir operated independently of any other reservoir. The complex stochastic problems in designing a system of dependent reservoirs are not dealt with.

The basic storage equation in design of a reservoir

$$I - O = \Delta S , \tag{1-9}$$

where I = the input, O = the output, and ΔS = the change in reservoir storage for a given time unit.

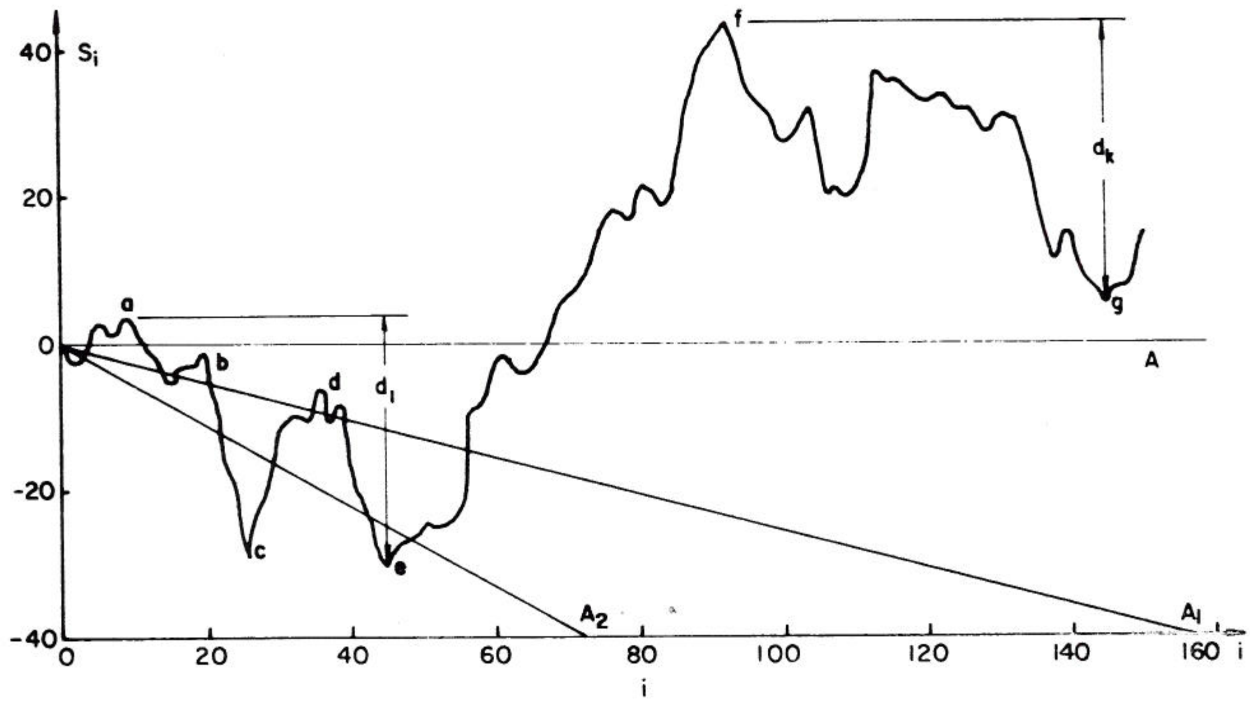

Fig. 1-3. Definition of Maximum Deficit, with Respect to Base OA Maximum Deficit = Max $\{d_i\}$, $1 \leq i \leq k$.

Neglecting the groundwater portion of the storage, and the seepage from a surface storage reservoir, but including the evaporation and the sedimentation of the reservoir, Eq. (1-9) can be rewritten as

$$X_t - Y_t - E_t = \frac{dS}{dt} \quad , \qquad (1\text{-}10)$$

where X_t = the input rate (such as daily discharge; Y_t = the output rate (such as daily water use); E_t = the evaporation rate from the reservoir, which depends mainly on the climate and the reservoir surface, and dS/dt = the rate of change in stored water.

Storage volume of a reservoir is a function of reservoir elevation and time. It can be approximated by

$$S = aH^m \qquad (1\text{-}11)$$

where $a = f(t)$ and $m = \phi(t)$ which are both functions of time because of reservoir sedimentation. The storage capacity, S_f, of a reservoir is always a finite value. It is a stochastic variable because $S_f = a(H^m_{max} - H^m_{min})$, with H_{max} and H_{min} the maximum and minimum reservoir water elevations, and a and m are stochastic variables. The evaporation is usually neglected in practical applications, when the average annual reservoir evaporation is small in comparison with the average annual input and output, and a and m are constants when the sediment inflow is small in comparison with the finite storage capacity. Accordingly Eq. (1-10) is modified as

$$X_t - Y_t = \frac{dS}{dt} \quad . \qquad (1\text{-}12)$$

It is *assumed* that short-interval flows, such as the daily river flows containing both periodic and stochastic components, are inputs into the reservoir. It is further assumed that there is no "trend" in the data, i.e., that man-made or natural changes in the river basin which produce these flows do not create significant trends in parameters. Thus the time series $x_{p,\tau}$ can schematically be presented as

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \, \varepsilon_{p,\tau} \qquad (1\text{-}13)$$

where $\tau = 1, 2, \ldots\ldots, \omega$, with ω the annual cycle (e.g., 12 months, 52 weeks, or 365 days), $p = 1, 2, \ldots\ldots, T$, with T = the number of years of record, μ_τ and σ_τ = the periodic mean and periodic standard deviation, respectively, and $\varepsilon_{p,\tau}$ = the stationary stochastic component with zero expectation and unit variance.

Periodic components can be described by harmonic functions. The stochastic component is usually assumed to follow the Markov linear models (first, second or higher order). The periodic component is the cyclic oscillation of means μ_τ (for each interval of the year, the mean is obtained over all years) and the cyclic component of standard deviation σ_τ for interval flows about the corresponding value of μ_τ.

The output is assumed to be composed of a stochastic component superimposed on periodic parameters. Assuming the trend component of the demand to be zero, though it does exist at least during the early period of operation of a reservoir until the stationary regime of storage operation is attained, it can be schematically presented by

$$Y_{p,\tau} = q_\tau + \delta_\tau f_{p,\tau} \quad , \qquad (1\text{-}14)$$

where τ and p have the same meaning as for the input process, while $Y_{p,\tau}$ represents the demand series, q_τ and δ_τ = the periodic (deterministic) mean and standard deviation respectively, and $\xi_{p,\tau}$ = the stochastic component.

Water releases from reservoirs are mostly assumed as deterministic processes. A rigorous mathematical description of outputs as stochastic processes is less feasible, when the outputs are regulated by reservoirs. Equation (1-14) is now modified to

$$Y_{p,\tau} = q_\tau \qquad (1\text{-}15)$$

Thus the net input to the reservoir is

$$x_{p,\tau} - Y_{p,\tau} = \mu_\tau - q_\tau + \sigma_\tau \, \varepsilon_{p,\tau} \qquad (1\text{-}16)$$

Now the regulation is called full, when

$$\sum_{\tau=1}^{\omega} q_\tau = \sum_{\tau=1}^{\omega} \mu_\tau \qquad (1\text{-}17)$$

Otherwise it is partial when

$$\sum_{\tau=1}^{\omega} q_\tau < \sum_{\tau=1}^{\omega} \mu_\tau \ , \qquad (1\text{-}18)$$

such that percentage of regulation or development p is equal to $\bar{q}_\tau/\bar{\mu}_\tau \times 100$.

The reservoir to be designed is assumed to be sufficiently *large* with a long economic life of more than 50 years.

It is further assumed that the demand is such that it corresponds to a *sufficiently high* level of development of the order of 90% or so. This assumption does not limit the application of this study to the practical problems because for large reservoirs the development is generally between 90-100%, and thus the assumption made is quite valid.

1-4 Objective of the Study

The objectives of this study are then the following:

(1) To determine the expected maximum deficit assumed to be the storage capacity needed for a reservoir under the following conditions: (a) a constant or a periodic deterministic output or demand, (b) large reservoir with long economic life, (c) partial regulation of river flow such that the level of development is higher than 90%, (d) short interval, e.g., daily flow, data for the input processes; and

(2) To determine the asymptotic distribution of the maximum deficit in case of higher-order Markov models for stochastic components.

1-5 Approach Used in this Study

Before going into the approach, a few descriptors of the periodic parameter ν_τ are first defined as

$$\bar{\nu}_\tau = \frac{\sum_{\tau=1}^{\omega} \nu_\tau}{\omega} \qquad (1\text{-}19)$$

and

$$s(\nu_\tau) = \sqrt{\frac{\sum_{\tau=1}^{\omega} (\nu_\tau - \bar{\nu}_\tau)^2}{\omega}} \ , \qquad (1\text{-}20)$$

where ν_τ represents both μ_τ and σ_τ and $\omega = 365$ when daily flow series are used.

In determining the storage capacity of a reservoir for within-the-year regulation with a deterministic demand, and for inputs of Markov models type with periodic mean and periodic standard deviation, the expected storage called herein as *total storage*, for the sake of brevity, is given by the expected maximum deficit of net input series $\{\mu_\tau + \sigma_\tau\varepsilon_{p,\tau} - q_\tau\}$. Similarly the expected maximum deficit of $\{\sigma_\tau\varepsilon_{p,\tau}\}$ is called *stochastic storage* while the difference between the total and the stochastic storages is called the *difference storage*.

If D_n^t and D_n^s are the random variables denoting the maximum deficit of net input $\{\mu_\tau - q_\tau + \sigma_\tau\varepsilon_{p,\tau}\}$ and stochastic net input $\{\sigma_\tau\varepsilon_{p,\tau}\}$ respectively, then

$$\text{Total Storage} = S_t = E[D_n^t]$$

$$\text{Stochastic Storage} = S_s = E[D_n^s]$$

Thus the total storage is divided into difference storage S_d and stochastic storage S_s written as

$$S_t(n) = S_s(n) + S_d(n) , \qquad (1\text{-}21)$$

where n is the number of time units.

Difference Storage. This part of the total storage is due to an extra term of $(\mu_\tau - q_\tau)$ in the net input series over and above the stochastic net input $\{\sigma_\tau\varepsilon_{p,\tau}\}$, and thus results from the difference between the periodic components of both input and output It is, in fact due to fluctuations of periodic components of input and output processes within the year. The difference storage is thus a function of periodic properties. It can be represented by

$$S_d(n) = f[\mu_\tau, \sigma_\tau, \rho, q_\tau, n] \qquad (1\text{-}22)$$

where $S_d(n)$ = the difference storage, ρ = the dependence of the stochastic component of the input process. To find the value of difference storage, the approach: difference storage = total storage - stochastic storage, will be used.

The point is that the difference storage, due to deterministic parts of input and output processes should become approximately constant by generating a *short* length of daily flow series. The difference storage is, therefore, found for different values of n (n being the sample size, or the number of time units), by finding the values of total storage and stochastic storage from the generated daily flow sequences. The process is repeated for a number of values of n until the value of difference storage becomes approximately constant within the sampling limits. This approximately constant value is called the *difference storage capacity*.

Stochastic Storage. This part of the total storage is a result of the difference of stochastic components of input and output processes. Since the stochastic component of the output is assumed zero, this storage is given by the expected maximum deficit of the stochastic component of the input process. It is, however, a function of σ_τ, and the dependence of the stochastic component. This can be schematically represented by

$$S_s(n) = f[\sigma_\tau, \rho, n] \quad , \tag{1-23}$$

where $S_s(n)$ = the stochastic storage, n = the subsample of the number of time intervals, ρ = the dependence of the stochastic component of the input.

Since the net stochastic input $\sigma_\tau \varepsilon_{p,\tau}$ of Eq. (1-13) is nonstationary, the stochastic storage is worked out first by finding the expected maximum deficit of a stationary process of $\varepsilon_{p,\tau}$, and then converting the same to give the stochastic storage for the nonstationary process of $\sigma_\tau \varepsilon_{p,\tau}$. Thus, the total storage can asymptotically be represented by

$$S_t(n) = S_d[\mu_\tau, \sigma_\tau, q_\tau, \rho] + S_s[\sigma_\tau, n, \rho] \quad ,$$

where $S_t(n)$ = the total storage or the needed storage capacity of a reservoir.

To determine the expected storage capacity or the total storage of a reservoir with long economic life the procedure is to find the value of difference storage capacity by generating a *short* length of daily flow series. The difference storage capacity, being independent of n for large values of n, is then taken as the same for the reservoir with the desired economic life. A coefficient C is also found from the above generated daily flow sequences which in turn is used in computing the value of stochastic storage for the given reservoir. The sum of the difference storage capacity and the stochastic storage thus gives the value of the total storage or the needed capacity of the reservoir. The advantage of the procedure is that it does not need generation of sequences equal to the economic life of the reservoir.

Asymptotic Distribution of Maximum Deficit of Markov Models. Gomide (1975) has proved analytically that for large values of n, the exact distribution of the *standardized* maximum deficit of partial sums of Markovian inputs tends to the asymptotic distribution of the *standardized* maximum deficit of the independent process. In case the asymptotic expected value and the variance of maximum deficit of higher-order Markov models are known somehow, then their asymptotic distribution can easily be worked out from the given distribution of the independent process. The effort in this study will be to find the asymptotic expected value and the variance of higher-order Markov models, and thereby to obtain the asymptotic distribution of maximum deficit for these models. The procedure holds good for any statistic of partial sums such as the surplus, deficit, and range.

Chapter 2
REVIEW OF LITERATURE

Reservoirs as a means of augmenting the low flows have been used for several thousand years. However, the first attempt to determine the size of a reservoir by a mathematical technique can be traced back only to the last century, when Rippl (1883) came up with his mass diagram. In spite of its limitations of not accounting for the stochastic nature of the input and output processes, the method still remains popular throughout the world. With the development of digital computers in the past 15 years, the experimental simulation or Monte-Carlo method combined with the Rippl's mass curve has been adopted for design of projects by studying alternate plans of operation.

Most of the literature review in this study can be divided broadly into two categories:

(1) Studies concerning infinite storage capacity; a great deal of research is on stationary processes studied by means of different statistics like surplus, deficit and range.

(2) Studies concerning finite reservoir size; these have been carried out by a few investigators. Most of these works relate to stationary processes, while actual hydrologic input and output processes of reservoirs are periodic-stochastic.

2-1 Analysis of Storage Problems by Range

W. Feller (1951) derived the asymptotic distribution of the range of the cumulative sums of independent normal random variables. In particular, he obtained the asymptotic mean and asymptotic variance of the range as

$$E(R_n) = 1.5958\sqrt{n} \qquad (2\text{-}1)$$

and

$$\mathrm{Var}\,(R_n) = 4n(\ell n\ 2 - \tfrac{2}{\pi}) \approx 0.2181\ n\ . \qquad (2\text{-}2)$$

He also found the expressions for the asymptotic mean and the asymptotic variance of adjusted range as

$$E(R_n^*) = \sqrt{n\pi/2} \approx 1.2533\sqrt{n} \qquad (2\text{-}3)$$

and

$$\mathrm{Var}\,(R_n^*) \approx \tfrac{\pi}{2}(\tfrac{\pi}{3} - 1)n \approx 0.0741\ n \qquad (2\text{-}4)$$

These theoretical results are independent of the underlying distribution of the original random variable having finite mean and finite variance, because for large values of n the partial sums S_n and S_n^* are asymptotically normally distributed.

A. A. Anis and E. H. Lloyd (1953) studied the problem of storage capacity of reservoirs for which the distribution of stored water is required over a number of years or of n time units. When the annual increments are independent variables with a common normal distribution, the water storage after i increments is the sum of i values. They gave the exact expected value of the maximum of the partial sums $S_1, S_2, S_3, S_4, \ldots\ldots, S_n$ of independent normal variables with mean zero and variance unity as

$$E(M_n) = \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{n} i^{-1/2} \qquad (2\text{-}5)$$

Because of symmetry the expected value of the range is

$$E(R_n) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1/2} \qquad (2\text{-}6)$$

Subsequently A. A. Anis (1955) gave the expression for the variance of the maximum of partial sums of a finite number of independent normal variates, which for $n \geq 2$ is

$$E(M_n^2) = \frac{1}{2(n+1)} + \frac{1}{2\pi} \sum_{i=1}^{n-2} \sum_{j=1}^{i} \{j(i-j+1)\}^{-1/2} \qquad (2\text{-}7)$$

with the asymptotic second moment of

$$E(M_n^2) = n - \frac{2+\sqrt{2}}{\pi}\, n^{1/2} \qquad (2\text{-}8)$$

A. A. Anis (1956) gave a recurrence relationship for obtaining the numerous evaluation of all the moments of partial sums of a finite number of independent normal variables.

A procedure for obtaining the exact distribution of M_n, m_n and R_n was described by Yevjevich (1965) for the values of $n = 2$ and $n = 3$. For higher values of n, Yevjevich used the data generation method to find the above distributions.

Using the data generation method, M. M. Melentijevich (1965) found the approximate equations for the expected value and variance of the range when the output is linearly dependent on storage.

V. Yevjevich (1967) suggested that the expected range of linearly dependent normal variables could be expressed by Eq. (2-9), which was derived for independent normal variables. He showed by data generation that the values calculated by Eq. (2-9) closely approximate the results obtained by data generation for the cases of first and second order autoregressive models and the simple moving average scheme.

$$E[R_n] = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} i^{-1}\, [\mathrm{Var}\ S_i]^{1/2}\ , \qquad (2\text{-}9)$$

where $\mathrm{Var}\ S_i$ is the variance of partial sum S_i. For independent normal variable $\mathrm{Var}\ S_i = i\sigma^2$. Notice that for $\sigma = 1$ Eq. (2-9) reduces to Eq. (2-6), as it should.

Francisco L. S. Gomide (1975) approached the problem of finding distributions of M_n, m_n and R_n by using Markov chains and found their numerical solution. He also found certain results for the first-order Markov inputs of stationary processes.

P. Sutabutra (1967) investigated the reservoir design problem for within-the-year flow regulation by separating the total storage into a deterministic storage part, which is a function of periodic means of inflow and outflow only, and a stochastic storage part, which is a function of stochasticities of input and output processes. The problem was studied under the following assumptions:

(1) Standard deviation of inflows at various time positions within the year is assumed constant;

(2) Stochastic component of monthly streamflow data follows the first-order Markov linear model;

(3) Stochastic component of the output is assumed to be nonexistent; and

(4) Output may be periodic, but the mean output is equal to the expected value of the input process.

Since no stochasticity in the output process is assumed, the stochastic storage in the form of the expected range for the first-order Markov model was determined.

Jose D. Salas-La Cruz (1972) followed the same approach as Sutabutra by relaxing only the first assumption of the constant standard deviation for inflows at various time positions within the year, because that assumption was not realistic. He in turn assumed a 12-month cycle in the standard deviation, thus making the results of his study applicable in designing reservoir capacities for monthly streamflow inputs. However, the applicability of his results was limited because of the following assumptions:

(1) Constant output was used instead of any deterministically changing output;

(2) 100% development and use of water resources rather than any partial use or development;

(3) The use of the first-order Markov model for the stochastic component of the input process (this may be true for monthly data but not for daily flow data in all cases, so that the method is not applicable to daily flow data); and

(4) The expected value of the range was used as the needed storage capacity of a reservoir rather than the expected maximum deficit.

2-2 Analysis of Finite Size Reservoir

P. A. P. Moran (1954) developed a simple formulation of the finite storage problem by making an extensive use of Markov chains, as a theory of storage with random inputs. Initially, a finite storage of a total capacity is available with independent inflows $X_t (t = 0, 1, 2, \ldots\ldots)$ in discrete units of time $(t, t+1)$. He presumed that a quantity Q_t already existed in storage before the arrival of the inflow X_t. Then if $X_t + Q_t > M$, some overflow will occur as the capacity of the reservoir is assumed to be M. The overflow is then $X_t + Q_t - M$, assuming either it is positive or zero. The storage contains a water quantity either M or $(X_t + Q_t)$, whichever is smaller. An amount Y may then be released according to some definite predetermined rule. A wide variety of such rules is possible. In one such rule Moran considered that a quantity of Y unit of water is released when

$$X_t + Q_t \geq Y \tag{2-10}$$

or a quantity $(X_t + Q_t)$ if this is less than Y. Continuing his work, Moran (1955) proceeded in several ways by considering different types of release rules.

C. H. Hardison (1965) dealt with the general problem of reservoir storage for low flow augmentation, but did not go deeply into any specific aspect of the problem. His main objective was to show as to how storage-draft relations can be related to probability so that developers can equate cost and risk more reliably. His approach to tackle the problem was to analyze the over-the-year storage separately from the analysis of within-the-year storage. He presented a method of combining the results by giving the probability of a given amount of storage required for selected draft rates. He claims that his method of finding storage-draft relationship is useful primarily in making preliminary estimates of potential development, in comparing the development possibilities of different streams, and comparing the alternate plans for development. Detailed design with variable draft rates would justify more sophisticated procedures like the Monte-Carlo mass curve method in which alternate plans of operation are simulated.

L. S. Gomide (1975) tackled the problem of partial regulation with theoretical analysis of mass curve. The storage problem concerning the partial regulation is called maximum accumulated deficit or maximum deficit analysis. This analysis was based on the application of Rippl's mass curve to the observed hydrologic sequences (Hurst, 1951), as shown in Fig. 1-3. This method is called by Fiering and some other authors as the "sequent-peak method" (Thomas and Fiering, 1963; Fiering, 1965).

Gomide (1975) derived the expression for the asymptotic expected maximum deficit for different inputs with mean zero and standard deviation unity as

$$E(D_n) = 1.2533 \sqrt{n} \tag{2-11}$$

Since Gomide considered only the stationary processes, his results can only be applied to annual inflows. Hence to design the reservoir for within-the-year regulation, the procedure is essentially the same as the old procedure by the Monte-Carlo method. In other words, sequences statistically indistinguishable from the actual record are simulated and then the same procedure is applied to each realization. Then probability distribution of maximum deficit is estimated by the relative frequency distribution. Usually the sample mean value is taken as the required storage (Fiering, 1965).

Brent M. Troutman (1976) found the limiting distribution of maximum deficit for the nonstationary dependent net input process of $g_\tau + \sigma_\tau \varepsilon_{p,\tau}$ where g_τ and σ_τ are the periodic mean and periodic standard deviation and $\varepsilon_{p,\tau}$ is the dependent or independent stochastic component of the net input process. He gave the limiting distribution for the case of full regulation when $\overline{g_\tau} = 0$ as

$$\lim_{n\to\infty} P[\frac{D_n}{\overline{\sigma}_\tau \sqrt{n}} \leq d] = F_D(d/\overline{\gamma}), \quad -\infty < d < \infty \tag{2-12}$$

where

$$\overline{\gamma}^2 = 1 + 2 \sum_{k=1}^{\infty} \overline{\rho}_k \ , \tag{2-13}$$

where $\bar{\rho}_k$ is the mean correlation coefficient at lag k.

In the case of full regulation the expectation of net input is zero, but when this value is positive, then he calls it "drift." The drift in fact corresponds to the case of partial regulation. In one case of drift he considers the expectation of net input as $c/\sqrt{n}$ which goes to zero as n goes to infinity as c is an arbitrary nonnegative constant. The maximum deficit in this case is denoted by $\tilde{D}_n$ and its distribution is given by

$$\lim_{n\to\infty} P[\frac{\tilde{D}_n}{\bar{\sigma}_\tau \sqrt{n}} \leq d] = F_D(d; \bar{\gamma}, \frac{c}{\bar{\sigma}_\tau}), \quad -\infty < d < \infty \quad (2\text{-}14)$$

He also considered a case of continuous drift when the expectation of net input does not converge to zero under assumptions A given as:

(1) Net input X_1 is i.i.d variable

(2) There exists a constant $p \neq 0$ such that

$$E[e^{-pX_1}] = 1 \quad ,$$

and

$$E[X_1 e^{-pX_1}] = -q > -\infty \quad (2\text{-}15)$$

and X_1 is nonlattice.

He found that the expected maximum deficit of i.i.d variables satisfying the above assumptions is a function of $\ell n(n)$.

A large number of references is available which contribute to the development of new approaches to the problem of storage in finite reservoirs, such as authored by N. U. Prabhu, A. Ghosal, G. F. Yeo, E. H. Lloyd, H. E. Hurst, and many others. It is unnecessary to cite all those who have contributed in one way or another to the theory of water storage to its present level. Only contributions directly related to this study have been discussed. However, for a comprehensive view the interested reader is referred to review papers by Lloyd (1967), Lloyd and Odoom (1964), Thomas and Fiering (1962) and Yevjevich (1964).

There are two characteristic features of most of these contributions which put them firmly in the domain of pure rather than of applied mathematics. The first of these features is the restriction to independent inflows (in the context of continuous time). The second feature which, while it contributed to mathematical elegance and tractability, has weakened the potentialities of the theory for engineering realism, was the tendency to abolish the top of the reservoir (Lloyd, 1974). Thus the practical problem of finding reliably the storage capacity of a reservoir with partial or full regulation still remained unresolved.

Chapter 3
DIFFERENCE STORAGE CAPACITY

3-1 Importance of Partial Water Flow Regulations

The basic consideration in designing storage capacities for irrigation, power generation, water supply and other uses, is to determine the reservoir size. In its simplest form the problem may be conceived as to determine the storage capacity required to secure a constant output equal to or close to the mean supply. Unfortunately, the problem is not simple in practice, where the percentage of development may not only be less than 100% but the output may also be periodic.

The fact is that the storage needed to guarantee a constant output, with the required output smaller than the input mean, may be more important from the particular point of view than when they are equal (Fathy and Shukry, 1956). Actually, the notion of constant output exactly equal to the input mean is not even practical due to the following:

(1) Input mean is not known in advance;

(2) To produce always a constant output equal to the input population mean would theoretically require an infinite storage capacity;

(3) Design based on the constant output equal to the input population mean would need significant part of storage capacity to be filled up at the start of operation to eliminate the empty reservoir states, with the amount of this initial storage difficult to determine beforehand. The flow regulation must necessarily deviate from an ideal program visualized for this constant output equal to input population mean; and

(4) A small reduction in the constant output below the input population mean leads to a relatively large reduction in the storage capacity needed. This allows for a reduction in the constant output, to be less than the input sample mean, because it serves as a factor of safety (because of the errors in estimated input mean) and as a way of securing a great economy in cost with a little sacrifice in benefit.

With these considerations in mind, the writer has decided to tackle this problem as a general case, namely of an output being smaller than the input population mean.

3-2 Definition of Difference Storage

As already mentioned in Section 1-3, short interval flows, such as daily flows, are the inputs to the reservoir in this study, represented mathematically by

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \varepsilon_{p,\tau} \quad . \qquad (3\text{-}1)$$

The daily output from the reservoir is represented mathematically by

$$Y_{p,\tau} = q_\tau \quad . \qquad (3\text{-}2)$$

The periodic components of input and output may be fitted by harmonic functions as given by Yevjevich (1972a). Their general form is

$$\nu_\tau = \bar{\nu} + \sum_{j=1}^{m} (A_j \cos\lambda_j \tau + B_j \sin\lambda_j \tau) \qquad (3\text{-}3)$$

where ν_τ = any periodic parameter, e.g., μ_τ, σ_τ, q_τ; $\bar{\nu}$ = the mean of the periodic parameter; $\lambda_j = 2\pi j/365$ for annual periodicity of daily values; m = the number of harmonics describing a periodic parameter; and p and τ as defined in Chapter I.

The net input into a reservoir is defined as the difference series $\{x_{p,\tau} - q_\tau\}$.

The stochastic part of input is $\sigma_\tau \varepsilon_{p,\tau}$, while the stochastic part of output is assumed zero. Hence, the stochastic part of net input into a reservoir is $\sigma_\tau \varepsilon_{p,\tau}$ given by $\{x_{p,\tau} - \mu_\tau\}$. It may be noted that the series $\sigma_\tau \varepsilon_{p,\tau}$ can be defined as input having a periodic standard deviation and zero mean. It could also be defined as the net input into a reservoir with periodic output being equal to the periodic input mean.

The *total storage capacity* needed for a sample size is defined as the *expected* maximum deficit of the net input. The distribution of this maximum deficit is estimated by generating a number *m* of net input series for a given sample of size n, and finding the frequency distribution of the maximum deficit. The total storage is then estimated by the mean maximum deficit of this frequency distribution. The maximum deficit for each sample is obtained by a sequent-peak algorithm.

The *stochastic storage capacity* is defined analogously as the *expected* maximum deficit of the stochastic net input. It is estimated by generating a number *m* of sample size n of stochastic net input and then finding the mean maximum deficit.

The *difference storage* is defined in this paper as the value of storage obtained by subtracting the expected maximum deficit of the stochastic net input from the expected maximum deficit of the net input, or

$$S_d(n) = S_t(n) - S_s(n), \qquad (3\text{-}4)$$

where

$$S_t(n) = E[D_n^t] \quad ,$$

$$S_s(n) = E[D_n^s] \quad ,$$

with $S_d(n)$ = the difference storage, $S_t(n)$ = the total storage, and $S_s(n)$ = the stochastic storage, the total storage is estimated by the mean of maximum deficit D_n^t of the net input $\{\mu_\tau - q_\tau + \sigma_\tau \varepsilon_{p,\tau}\}$ series, while the stochastic storage is estimated by the mean of maximum deficit D_n^s of the stochastic net input $\{\sigma_\tau \varepsilon_{p,\tau}\}$ series only. The difference between these two is the estimate of difference storage. It is in fact the storage required due to an additional term $(\mu_\tau - q_\tau)$ in the net input series. Since this term is deterministic, the difference storage required may be approximately constant for large values of n.

Since the total storage of Eq. (3-4) is a function of the periodicity of input and output, the difference storage is a function of harmonics of means of water supply and water demand, or of their frequencies, amplitudes, and phases.

3-3 Aim of this Chapter

The aim in this chapter is to investigate the behaviour of difference storage for different values of n with the help of computer simulation of a periodic-stochastic input and of any deterministically changing output. It will be shown from computer results that asymptotically the value of difference storage is approximately constant within the sampling variation, at least for a sufficiently high level of regulation. This can be estimated by generating a short length of daily flow series. This value of the difference storage, approximately constant, called as *difference storage capacity*, can in turn be used for the design of a reservoir with a long economic life.

3-4 Model Used for Input Process

To study the behavior of difference storage for various values of n, the periodic-stochastic input or daily discharge series for selected series lengths are generated by using the *same* tape of 250,000 standard normal random numbers rather than generating new random numbers for different runs, resulting in some common random numbers in some series. With the assumed given output, the difference storage is then estimated for various values of n after the values of total and stochastic storages have been found.

The periodic-stochastic process is generated (Yevjevich, 1966) by using

$$x_{p,\tau} = \hat{\mu}_\tau + \hat{\sigma}_\tau(\bar{y} + s_y \varepsilon_{p,\tau}) \quad , \qquad (3\text{-}5)$$

with $\bar{y}$ and s_y = the mean and standard deviation of $(x_{p,\tau} - \hat{\mu}_\tau)/\hat{\sigma}_\tau$. $\varepsilon_{p,\tau}$ = the dependent or independent stationary stochastic component with mean zero and standard deviation unity. Quimpo (1967) has inferred that the stochastic component of daily flows approximately follows the second-order Markov model. $\varepsilon_{p,\tau}$ is, therefore, represented as

$$\varepsilon_{p,\tau} = a_1\varepsilon_{p,\tau-1} + a_2\varepsilon_{p,\tau-2} + s\eta_{p,\tau} \quad , \qquad (3\text{-}6)$$

where

$$s = [1 - a_1^2 - a_2^2 - \frac{2\, a_1^2\, a_2}{(1 - a_2)}]^{1/2} \quad , \qquad (3\text{-}7)$$

$\eta_{p,\tau}$ = the independent standard random numbers, assumed as normal in this investigation, where the interest is only for asymptotic results as outlined in Chapter IV. Its distribution is not important because for large values of n, the partial sums are asymptotically normally distributed.

To generate the daily flows by Eq. (3-5) the following values of $\bar{y}$, s_y, and $\hat{\mu}_\tau$ are assumed:

$\bar{y}$ = 0.034, s_y = 1.174, a_1 = 0.5418, a_2 = 0.3193, and

$$\hat{\mu}_\tau = 543.498 + \sum_{j=1}^{m} (A_{mj}\cos \lambda_j\tau + B_{mj}\sin \lambda_j\tau) \qquad (3\text{-}8)$$

$$\hat{\sigma}_\tau = 288.370 + \sum_{j=1}^{m} (A_{sj}\cos \lambda_j\tau + B_{sj}\sin \lambda_j\tau)$$

The values of A_m, B_m, A_s, and B_s are given for various harmonics in Table 3-1.

Table 3-1. Fourier Coefficients of Periodic Mean and Periodic Standard Deviation used for Generating the Input Process

Harmonic	Fourier Coefficients of Periodic Mean		Fourier Coefficients of Periodic Standard Deviation	
	A_m	B_m	A_s	B_s
1	- 200.30	- 112.40	- 123.30	- 85.60
2	145.40	185.00	141.60	105.70
3	- 85.50	- 79.90	- 66.40	- 46.20
4	58.00	65.60	75.70	31.70
5	- 39.80	- 72.50	- 47.20	- 43.20

3-5 Characteristics of Total Storage Capacity

The properties of the total storage affecting the computation of difference storage are first discussed.

Rate of Increase of Total Storage Capacity of Periodic-Stochastic Process in Case of Different Percentages of Development. The rate of increase of total storage capacity with n depends on the desired percentage of development. The study has been carried out on generated samples of daily flows by Eq. (3-5). The expected maximum deficits for 100% and 90% developments are tabulated in Table 3-2 and plotted in Figure 3-1, as an illustration of the effects of the level of development. It is clear that the rate of change of total storage increases rapidly as the level of development approaches the input population mean. This is also an intuitive conclusion, because for zero output no reservoir is required, with the maximum deficit zero for every n. It may be pointed out that, first the two curves in Fig. 3-1 do not seem to have the same asymptotes, and second, the confidence bands of the two curves widen with an increase of n, because of the reduction in the number of generated series from the same 250,000 random standard normal numbers.

Table 3-2. Total Storage Capacity for Different Constant Outputs for Periodic-Stochastic Process

Sample No.	n in days	Type of Output: 100% Development	90% Development
1	365	27637.901	20113.064
2	730	43237.297	29076.202
3	1095	52228.577	34332.548
4	1460	58408.543	37432.759
5	1600	63472.517	39742.526
6	1825	64837.067	40867.569
7	2190	68806.653	42919.674
8	2920	74232.883	45310.545
9	3200	80364.674	47280.794
10	3650	81963.895	47151.410

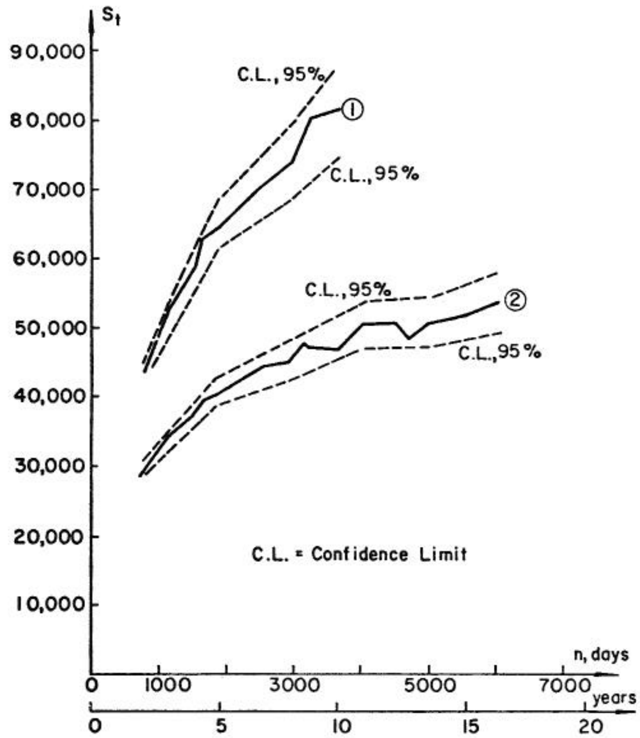

Fig. 3-1. Total Storage Capacity Versus Development Level for Periodic-Stochastic Processes: (1) 100% Development, (2) 90% Development.

Effect of Phase Difference in Input and Output on Rate of Increase of Total Storage Capacity of Periodic-Stochastic Processes. Input is generated by Eq. (3-5) and output is assumed as

$$q_\tau = 515.6 + A_1\cos\lambda_1\tau + A_2\cos\lambda_2\tau + B_1\sin\lambda_2\tau + B_2\sin\lambda_2\tau \quad , \tag{3-9}$$

with $\lambda_j = 2\pi j/365$. Two cases are studied. The first case is: $A_1 = -344.00$, $B_1 = +147.80$, $A_2 = 79.00$, and $B_2 = -46.10$, so that the phase difference of the first harmonic in input and output is approximately 53°. The output is shown in Fig. 3-2.

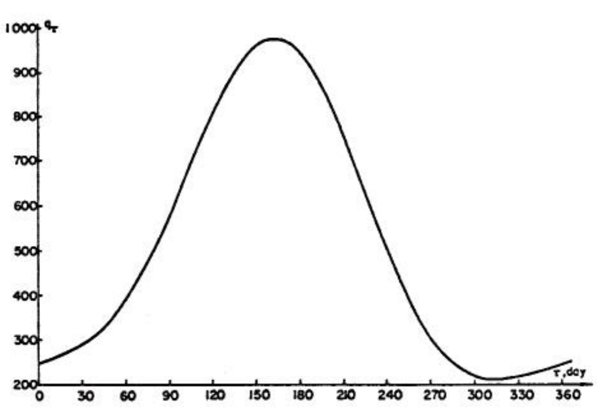

Fig. 3-2. Values of Periodic Output Within the Year with $\bar{q}_\tau = 515.6$, and $(\phi_1 - \delta_1) = 53°$ approx.

The second case is: $A_1 = -187.21$, $A_2 = 79.00$, $B_1 = 324.24$, and $B_2 = -46.10$, so that the phase difference of first harmonics of input and output is 90°, while the amplitudes of harmonics are unchanged.

Total storage found for different n values for these two cases of differences in phases, are plotted in Fig. 3-3. Though the curves (1) and (2) are almost parallel, the rate of increase of total storage capacity as function of n does not depend on phase differences of input and output harmonics, so long as the mean output is the same.

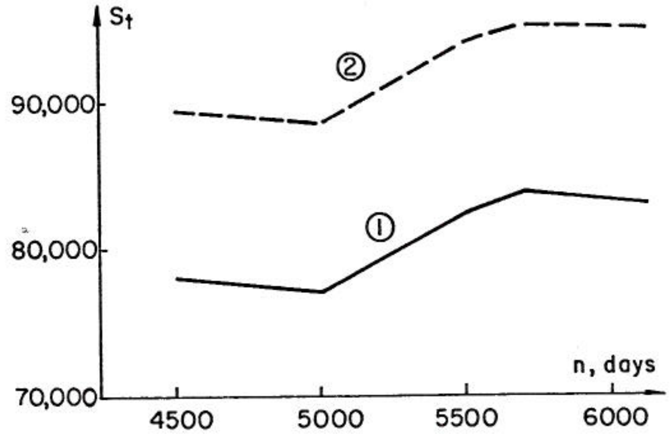

Fig. 3-3. Total Storage Capacity Versus Phase Difference for Periodic-Stochastic Processes: (1) Phase Difference $\phi_1 - \delta_1 = 53°$; and (2) Phase Difference $\phi_1 - \delta_1 = 90°$.

Effect of Phase Difference on the Total Storage Capacity. Figure 3-3 shows that for the same mean output the expected maximum deficit or the total storage capacity increases as the phase difference increases from 53° to 90°.

3-6 Characteristics of Stochastic Storage Capacity

Effect of Output on Stochastic Storage Capacity. As discussed in Section 3-2, the stochastic net input to a reservoir is defined by $\{x_{p,\tau} - \mu_\tau\}$. It remains the same as long as the output is deterministic. Hence, the stochastic storage determined as the expected maximum deficit of this series remains the same for all deterministic outputs.

3-7 Characteristics of Difference Storage Capacity

Asymptotic Value of Difference Storage as a Constant. It is reasonable to conjecture that since the difference storage is due to an extra deterministic term $(\mu_\tau - q_\tau)$ in the net input series over and above the stochastic net input series, and as such is not a result of any stochasticity, its value should at least be asymptotically constant. It may not be a constant for small values of n because the difference storage is defined by Eq. (3-4), wherein both the total and the stochastic storages are functions of n. Hence the difference storage would become constant only when the rate of increase of the total storage and the stochastic storage with n is the same. Since the rate of increase of the total storage depends on the level of development, as shown in

Section 3-5, while the rate of increase of the stochastic storage is independent of any deterministic output, as shown in Section 3-6, the asymptotic behavior of difference storage is studied separately for the two cases of the level of development.

Case I. Full Development. In this case both the total and the stochastic storages correspond to the same 100% development with the difference that the former has some phase difference and amplitude ratio, while the latter has a zero phase difference with an amplitude ratio of one. Assuming an amplitude ratio of one, both for the total and the stochastic storages, the difference storage, defined as the difference between total and stochastic storage, would just be the storage required due to an effect of phase difference in the net input series.

The effect of changing the phase difference on the value of total storage was studied for the general case of partial development. The results so obtained from computer runs are already plotted in Fig. 3-3. It is found that, for the *same* percentage of development or for the *same* annual mean output, the rate of increase of the total storage with n is independent of the phase difference between the periodic means of input and output.

This is quite intuitive because the change in phase difference is responsible only for the fluctuations of the net input process within the year. Hence it must result in the change of value of the total storage. So far as the stochasticity of the net input is concerned, it is not affected in any way. Therefore the rate of increase of the total storage with n should not be altered by the change in the phase difference for the *same* level of development.

Thus, taking the two special cases of 100% development or the full regulation of Fig. 3-3, curve (1), may correspond to an amplitude ratio of one and a phase difference of zero, i.e., q_τ is equal to μ_τ at all times within the year, while curve (2) may correspond to the full regulation with an amplitude ratio of one and a phase difference of more than zero. According to the definitions, curve (2) will give the total storage for this particular case, while curve (1) will give the stochastic storage. It is then clear from Fig. 3-3 that the difference between curves (2) and (1), being the value of the difference storage, is a constant within the sampling variations for this range of n values.

This computer result can be supported by considering the special cases of the general result given by Troutman (1976). He considered a net input process as

$$W_{p,\tau} = g_\tau + \sigma_\tau \varepsilon_{p,\tau} \quad , \qquad (3\text{-}10)$$

where $W_{p,\tau}$ = net input and g_τ and σ_τ are the periodic mean and periodic standard deviation of the net input process. Comparing this with the net input process of this study, it is found that

$$g_\tau = \mu_\tau - q_\tau \quad . \qquad (3\text{-}11)$$

He gave the limiting distribution of maximum deficit for the case of full regulation when $\bar{g}_\tau = 0$ as

$$\lim_{n\to\infty} P[D_n^t / \bar{\sigma}_\tau n^{1/2} < d] = F_D(d/\bar{\gamma}) \quad , \qquad (3\text{-}12)$$

where $\bar{\gamma}$ is defined by *his* Eq. (4-47). From Eq. (3-12) the expectation of maximum deficit can be written as

$$E[D_n^t] \simeq c_1 \bar{\sigma}_\tau n^{1/2} + o(n^{1/2}) \quad , \qquad (3\text{-}13)$$

where c_1 is constant and the second term $o(n^{1/2})$, if constant, also satisfies Eq. (3-12) supporting the evidence of constant value obtained from Fig. 3-3.

Let us now consider two special cases of the general result of Eq. (3-12). The first case is for an amplitude of one and a phase difference of more than zero. In this case Eq. (3-13) can be modified as per definition of total storage as

$$S_t \simeq c_1 \bar{\sigma}_\tau n^{1/2} + d_1 \quad , \qquad (3\text{-}14)$$

where d_1 is a constant. The second case is for an amplitude of one and a phase difference of zero. In this case Eq. (3-13) can be modified as per definition of stochastic storage as

$$S_s \simeq c_1 \bar{\sigma}_\tau n^{1/2} + d_2 \quad , \qquad (3\text{-}15)$$

where d_2 is also a constant. Since both total and stochastic storages have asymptotically similar functions of $\sqrt{n}$ in Eq. (3-14) and (3-15), their difference, defined as difference storage, can be asymptotically constant.

To test the validity of this asymptotically constant value of difference storage, the daily flows are simulated by Eq. (3-5), and an output corresponding to 100% development is considered in order to compute the total storage and the stochastic storage of a periodic-stochastic process. The difference storage is then obtained as the difference of these two storage capacities. The results so obtained are given in Table 3-3 and plotted in Fig. 3-4.

From Fig. 3-4 it is seen that the difference storage first increases with an increase of n and then stabilizes to an approximately constant value of 16270, which is obtained by drawing an average line on Fig. 3-4 after the difference storage starts oscillating about this line. The increase with n occurs because the rate of increase of total storage for 100% development is higher than the rate of increase of stochastic storage until the two rates of increase with n become the same, to give a constant value of difference storage within sampling error limits, thereby supporting the conclusion arrived above regarding the asymptotically constant value of difference storage.

If the hypothesis of expressing S_t and S_s by Eqs. (3-14) and (3-15) is true, then the variances of maximum deficit of net input and that of stochastic net input must be the same for each n in the region where the difference storage gets stabilized to a nearly constant value. Hence, the variances of D_n^t and D_n^s are calculated from the computer results and given in Table 3-4.

It will be seen from Table 3-4 that the variances of maximum deficits of net input and stochastic net input are nearly the same. Moreover, since the vari-

Table 3-3. Difference Storage for Various Values of n for Constant Outputs Corresponding to 100% and 90% Developments

n in days	100% Regulation: Total Storage	100% Regulation: Stochastic Storage	100% Regulation: Difference Storage	90% Regulation: Total Storage	90% Regulation: Stochastic Storage	90% Regulation: Difference Storage
365	27637.901	20894.261	6743.640	20113.064		- 781.197
730	43237.297	31775.757	11461.540	29076.202		- 2699.555
1095	52228.577	38432.445	13796.132	34332.548		- 4099.898
1400	58402.543	43293.396	15115.147	37432.759		- 5860.637
1600	63472.517	46404.470	17068.047	39742.526		- 6661.945
1825	64837.067	48601.479	16235.588	40867.569	Same as for	- 7733.910
2190	68806.653	54253.030	14553.623	42919.674	100%	- 11333.356
2920	74232.883	57103.457	17129.426	45310.545	Development	- 11792.912
3200	80364.674	64692.777	15671.897	47280.794		- 17411.983
3650	81963.895	65005.850	16958.045	47151.410		- 17854.440
4000				50356.722	68832.532	- 18475.810
4200				50692.421	68796.583	- 18104.162
4500				50834.850	73403.330	- 22568.480
4700				48742.237	72133.919	- 23391.683
5000				50947.593	73627.779	- 22680.180
5500				51960.668	80884.754	- 28924.080
6000				53914.483	77517.978	- 23603.490
6500				54492.398	82766.646	- 28274.240
7000				55011.934	79719.656	- 24707.720

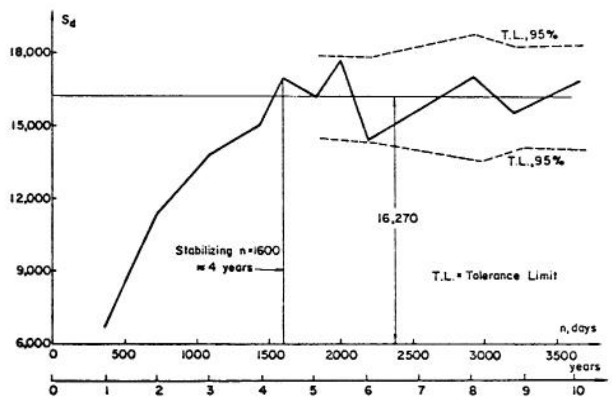

Fig. 3-4. Difference Storage Versus n for Constant Output Corresponding to 100% Development.

Table 3-4. Variance of Maximum Deficits of Net Input and Stochastic Net Input for Different Values of n for Constant Output Corresponding to 100% Development

n in days	$Var[D_n^t]$ $x10^8$	$Var[D_n^s]$ $x10^8$
1825	4.05	4.56
2920	6.51	6.32
3650	7.91	7.94

ance of D_n^t grows as a function of n, being almost twice for n = 3650 than that of n = 1825 in Table 3-4, it further supports the hypothesis of expressing S_t and S_s by Eqs. (3-14) and (3-15), which in turn supports the conclusion of asymptotically constant value of difference storage.

Case II. Partial Development. It has already been concluded that in case of full regulation the changes in phases etc., of periodic means of net input result in the same rate of increase of total and stochastic storages, as these changes do not alter the stochasticity of the net input process. Since in the case of partial development with $\bar{q}_\tau$ less than $\bar{\mu}_\tau$, the stochasticity of the net input remains the same, the expectation is that *asymtotically* the total and stochastic storages *may* still grow with the same function of $\sqrt{n}$. If this is true, then the asymptotic value of the difference storage would be a constant reflecting the properties of the periodic component of the net input process.

Troutman (1976) studied the problem of partial regulation for the case of i.i.d net input random variables, X_i, under the assumption that there exists a constant $p \neq 0$ such that

$$E[e^{-pX_i}] = 1$$

and

$$E[X_i e^{-pX_i}] = -q > -\infty$$

where X_i is non-lattice. For example, this assumption is satisfied if X_i is normally distributed with the mean μ and the variance σ^2. In this case $p = 2\mu/\sigma^2$ and $q = \mu$.

These assumptions were made to bring the problem into the framework of random walk theory. Using the random walk approach he found that the expected maximum deficit for the net input process of i.i.d variables, X_i, grows asymptotically as a function of $\ell n\ n$ rather than the $\sqrt{n}$ corresponding to the case of full regulation. Since periodicity did not change the asymptotic behavior for full regulation it may not change the asymptotic behavior here. Hence, it is not clear whether or not the expected maximum deficit of nonstationary independent process should grow as a function of ℓn n.

To answer this question the daily flow series is simulated by Eq. (3-5) with the stationary stochastic component $\varepsilon_{p,\tau}$ as an independent standard normal random variable, and a constant output corresponding to 90% development is considered to compute the total storage and the stochastic storage of a periodic-stochastic process. The difference storage is then obtained as the difference of these two storage capacities. The results so obtained are given in Fig. 3-5.

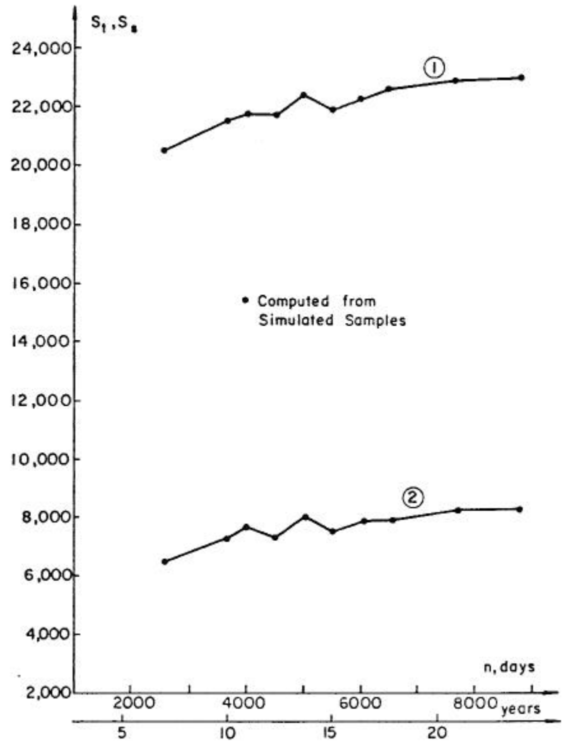

Fig. 3-5. Total Storage and Stochastic Storage for Periodic-Stochastic Independent Process with Constant Output Corresponding to 90% Development: (1) Total Storage; and (2) Stochastic Storage.

From Fig. 3-5 it is found that curves (1) and (2) corresponding to the total storage for partial development and the stochastic storage (defined as a case of full development with $q_\tau = \mu_\tau$), are almost similar in shape, indicating thereby that total storage for partial development may also grow with the same function of n as that of stochastic storage for large values of n (the values of n used in Fig. 3-5 are of the order of 15-23 years). Since stochastic storage grows as a function of $\sqrt{n}$ so the total storage may also grow as a function of $\sqrt{n}$, so that the asymptotic rate of increase of total storage is independent of the level of development, which may be due to no alteration in the stochasticity of the net input process by changing the level of development.

To further investigate the growth function of total storage for partial development, two regression lines are fitted to curve (1) of the total storage: both as a function of $\sqrt{n}$; and as a function of ℓn n. The two regression lines are

$$S_t = 55.35 \sqrt{n} + 18047.25 \quad , \qquad (3\text{-}16)$$

and

$$S_t = 1981.72 \ \ell n \ n + 5142.14 \quad . \qquad (3\text{-}17)$$

The two lines explain almost the same percentage of variance, the former explaining 91% while the latter 93%. The two regression lines are plotted in Fig. 3-6 as curves (2) and (3). It is found that the two curves are almost the same in the region of study. The regression line fitted for the stochastic storage is

$$S_s = 36.55 \sqrt{n} + 5025.53 \quad . \qquad (3\text{-}18)$$

It explains 80% of the variance and is shown as curve (1) in Fig. 3-6.

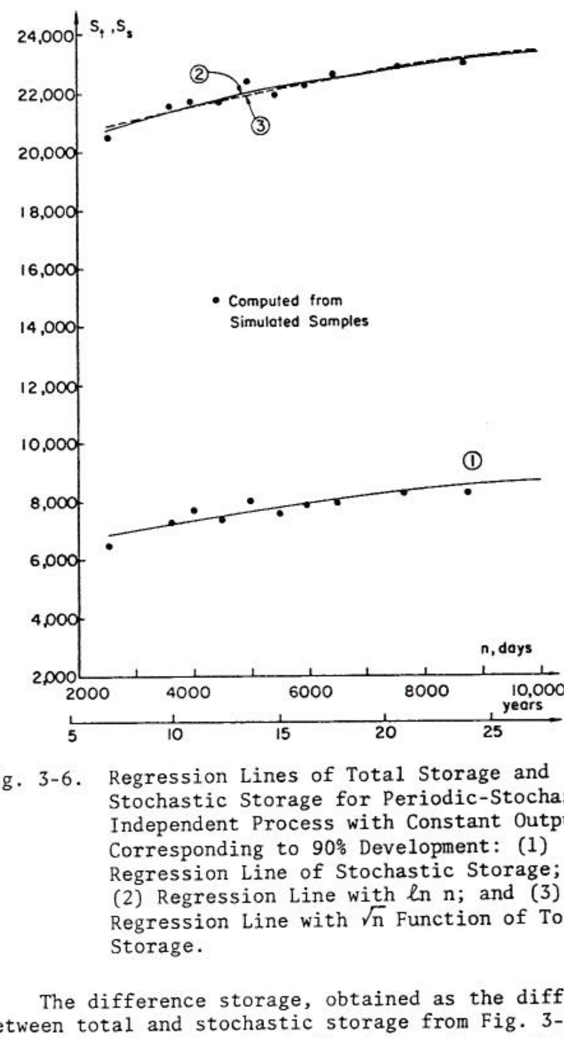

Fig. 3-6. Regression Lines of Total Storage and Stochastic Storage for Periodic-Stochastic Independent Process with Constant Output Corresponding to 90% Development: (1) Regression Line of Stochastic Storage; (2) Regression Line with ℓn n; and (3) Regression Line with $\sqrt{n}$ Function of Total Storage.

The difference storage, obtained as the difference between total and stochastic storage from Fig. 3-6, is plotted in Fig. 3-7. It is found from Fig. 3-7 that the difference storage capacity is approximately constant and is equal to 14,400 within the errors of sampling variations.

The above result may be because of the repeated use of the same random numbers for different values of n. Two values of difference storage corresponding to n = 2555 and n = 8760 were obtained by generating *different* random numbers directly by the computer.

Even then the values of difference storage are of the order of 14,400 within sampling variations. However, this stabilization of difference storage, shown by curve (2) in Fig. 3-7, to an approximately constant value could result from the fact that curves (2) and (3) in Fig. 3-6 are almost the same in the region of study.

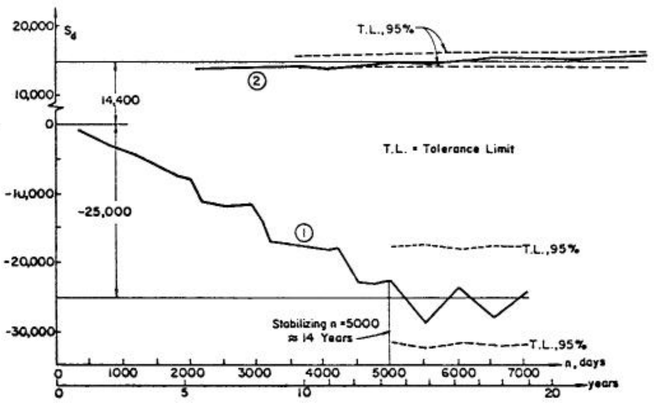

Fig. 3-7. Difference Storage Versus n for Constant Output Corresponding to 90% Development: (1) Difference Storage for Dependent Process; and (2) Difference Storage for Independent Process.

It is expected that unless the total storage grows with the same function of $\sqrt{n}$ as the stochastic storage, the variances of D_n^t and D_n^s would not be the same function of n in the region where the difference storage is found to be approximately constant in Fig. 3-7. The variances of D_n^t and D_n^s are, therefore, computed from the computer results and given in Table 3-5. The computed variances do not seem to grow as the theory dictates, which may be because of their not being very precise estimates.

Table 3-5. Variances of Maximum Deficits of Net Input and Stochastic Net Input for Different Values of n for Periodic-Stochastic Independent Input and Constant Output Corresponding to 90% Development

n in days	$Var[D_n^t]$ $x10^6$	$Var[D_n^s]$ $x10^6$
4500	3.78	3.41
5500	3.55	3.73
6500	3.18	2.91
7665	2.90	3.02
8760	3.93	3.33

It is found from Table 3-5 that the variances of maximum deficits of net input and stochastic net input are nearly the same for large values of n, for which the difference storage is approximately constant, thereby supporting the contention that *asymptotically* the total storage may grow as a function of $\sqrt{n}$ in the case of partial development too. This may be so, because the assumptions made by Troutman (1976) to bring the problem in the frame work of random walk approach are true for his case of stationary independent net input process. But the assumptions are certainly not true for the nonstationary independent process being studied herein, because in this case the reservoir level does not follow a random walk but in turn either goes on increasing during the flood period or goes on falling during the recession of the flood. However, since the computer results cannot prove or disprove a theory, it is felt that further theoretical work must be undertaken to determine whether the total storage or expected maximum deficit for the periodic-stochastic input and the deterministic output grows as a function of $\sqrt{n}$ or ℓn n, or some other function for the case of partial development.

The computer results of this analysis, however, lead to the conclusion that the difference storage becomes approximately constant within the region of study, corresponding to large values of n at least in case of a *high* level of development, namely of the order of 90% or so.

Effect of Dependence in Stochastic Variable of Input Process. The daily flows are simulated by Eq. (3-5), and an output corresponding to 90% development is considered to compute the total storage and stochastic storage of periodic-stochastic process. The results so obtained are given in Table 3-3 and plotted in Fig. 3-7. It can be inferred from Table 3-3 and Fig. 3-5 that the dependence in stochastic variable increases both the total and the stochastic storage, but the increase in the stochastic storage is greater than in the total storage. Hence, the effect of introducing the dependence in the stochastic variable of the input process is to reduce the asymptotic value of the difference storage. In this particular case this value or the difference storage capacity gets reduced from 14,400 to (-) 25000.

Possibility of Negative Value of Difference Storage Capacity. It can be noted from Table 3-3 that for 90% development the difference storage is negative. It is justified because of the fact that the total storage needed for the reservoir depends on the output, while the stochastic storage is independent of output as per Section 3-6. Hence, as the output is reduced the total storage is reduced, while the stochastic storage remains the same. Thus a limit can be conceived when the output is too low such that the total storage required is less than the stochastic storage. The difference storage, being the difference between these two, would then be negative.

Intuitively it can also be conceived. As already explained in Section 3-2, the difference storage is the storage required because of the additional term $(\mu_\tau - q_\tau)$ in the net input series over and above the stochastic net input series. If q_τ is assumed to be equal to or smaller than the least value of μ_τ during the year, then there will be no requirement of storage due to $(\mu_\tau - q_\tau)$. Rather an extra supply of μ_τ would reduce the storage requirement of the stochastic component of the input process, thereby explaining the possibility of negative difference storage. This explains the feasibility of conceiving the negative difference storage capacity.

Percentage of Development. It is obvious that a lower percentage of development would require a lower capacity of the reservoir. But since the stochastic storage is independent of the percentage of development, the reduction in the required storage capacity should be affected by the corresponding reduction in the value of the difference storage. Referring to

Figs. 3-4 and 3-7, it may be noticed that there is a drastic reduction in the value of the approximately constant difference storage from 16,270 to (-) 25000, as the percentage of development is reduced from 100% to 90%, a fact recognized by Hurst (1951), who observed that a small reduction in the guaranteed output from the maximum value (the mean) makes a great proportional reduction in the storage required to maintain it.

Effect of Phase Difference of Input and Output. The phase lags between the input and output harmonics have a bearing on the total storage and consequently on the value of the difference storage, as the stochastic storage is independent of the phases of periodic components of input and output. To save computer time only two harmonics in the periodic mean of input are taken in Eq. (3-8) instead of five. The phase lag of the first harmonic of input and output was varied, and the values of difference storage calculated are given in Table 3-6 and plotted in Fig. 3-8. It is clear from Fig. 3-8 that the difference storage is a function of the phase difference between input and output.

Effect of Amplitude of Output. The study has been conducted similar to above except that the amplitude of output is varied keeping the same phase difference between the first harmonic in periodic means of input and output. The results are given in Table 3-6. It shows that even for the same phase difference between input and output the difference storage increases directly with the increase of the amplitude ratio.

By studying various factors affecting the difference storage, it is concluded that *asymptotically* its function is

$$S_d = f[\mu_\tau, \sigma_\tau, p, \psi_i, \rho, r_i] \quad , \qquad (3\text{-}19)$$

where S_d = the asymptotic value of difference storage, p = the percentage of development, ψ_i = the phase difference of the i-th harmonic of input and output processes, and r_i = the ratio of amplitudes of output and input harmonics.

3-8 Stabilization Region for Difference Storage Capacity

So far it has been found from computer results that for large values of n the difference storage is approximately constant. However, the question arises, after what values of n the difference storage would stabilize to an approximately constant value so that the same may be adopted in the design of reservoir with a long economic life beyond this value of n. Figs. 3-4 and 3-7 show that the asymptotic value of difference storage can be found by simulation. But there is always a transition region wherein the difference storage either increases or decreases continuously till it attains an approximately constant value. This transition region is called here as the stabilization region. In Fig. 3-4 this region is for $n < 1600$ while in Fig. 3-7 it is for $n < 5000$. The length of stabilization region thus determines the minimum length of a series to be simulated in using this method of determining the asymptotic value of difference storage. Hence, the longer the stabilization region, the more is the effort needed in finding this value.

Jose D. Salas-La Cruz (1972) analyzed the above problem with a different definition of difference storage. His analogous term is deterministic storage, which he defined as the difference of the expected range of net input series and the expected range of stochastic net input series. Though his definition of deterministic storage is different from the definition used in this study, there is an analogy between his results with those obtained in this study. He found through generation that in case of monthly periodicity of mean and standard deviation, there is a transition region where the influence of phases of μ_τ and σ_τ was significant, namely the region for $n < 50$. Beyond this value the deterministic storage becomes approximately constant.

It is the conviction of this writer that the transition region, or the stabilization region as called in this study, is not only due to the influence of phases of μ_τ and σ_τ, but it also depends on the level of development. In fact, it mainly depends on the rate of increase of total storage vis-a-vis the rate of increase of the stochastic storage with n. As explained in Section 3-7, the difference storage would become constant only when the total and the stochastic storages have the same rate of increase with n. Since the rate of increase of total storage changes with the percentage of development, being higher with higher development and lower with lower development, the stabilization region would be longer for partial regulation than for full regulation. This is shown heuristically below.

Derivation of Stabilization Region. The heuristic condition for the stabilization of difference storage to an approximately constant value can be written as

$$\frac{dS_t}{dn} = \frac{dS_s}{dn} \quad . \qquad (3\text{-}20)$$

Consider two different percentages of development to find out the difference in the lengths of stabilization region.

Case I — 100% development

Let $S_t(p)$ and $S_d(p)$ denote the total and the difference storages for p % development and S_s is the stochastic storage such that

$$S_t(100) \simeq a\, n^c + S_d(100) \quad . \qquad (3\text{-}21)$$

$$S_s \simeq b\, n^d \quad , \qquad (3\text{-}22)$$

where a and b are proportionality constants and c and d are the powers with which the total and the stochastic storages vary with n as *approximations* to the true laws of variations of S_t and S_s with n. Then for the region of n where S_d is approximately constant

$$\frac{dS_t(100)}{dn} \simeq a\, c\, n^{c-1} \quad . \qquad (3\text{-}23)$$

$$\frac{dS_s}{dn} \simeq b\, d\, n^{d-1} \quad . \qquad (3\text{-}24)$$

Now the difference storage would attain approximately constant value only when the rate of increase of total storage and stochastic storage with n is the same. Let this point correspond to $n = n(100)$. Therefore, Eqs. (3-23) and (3-24) should be equal at $n = n(100)$, so that

Table 3-6. Difference Storage for Various Phase Differences

Ratio of Amplitudes / Phase Difference (φ-δ)	1.0			2.66			0.0		
	Total Storage	Stochastic Storage	Difference Storage	Total Storage	Stochastic Storage	Difference Storage	Total Storage	Stochastic Storage	Difference Storage
0°	34949.502	34949.502	0.000	42460.495	34949.502	7510.994	44356.935	34949.502	9407.433
45°	44594.600	"	9645.099	-	-	-	"	"	"
90°	55962.799	"	21013.297	69196.871	34949.502	34247.369	"	"	"
180°	67028.761	"	32079.259	82872.141	"	47922.639	"	"	"
270°	57903.536	"	22954.034	71724.346	"	36774.844	"	"	"
315°	44584.040	"	9634.547	54391.172	"	19441.670	"	"	"
360	34949.502	"	0.000	42460.495	"	7510.994	"	"	"

Note: (1) These values have been found for $n = 1460$ days;

(2) Periodic mean of input $\mu_\tau = 543.498 - 200.30 \cos \lambda_1\tau - 112.40 \sin \lambda_1\tau + 145.40 \cos \lambda_2\tau + 185.00 \sin \lambda_2\tau$;

(3) Periodic output, $q_\tau = 543.498 - A_1 \cos \lambda_1\tau - B_1 \sin \lambda_1\tau$; and

(4) A_1 and B_1 are varied to get different phase differences for the given amplitude ratio.

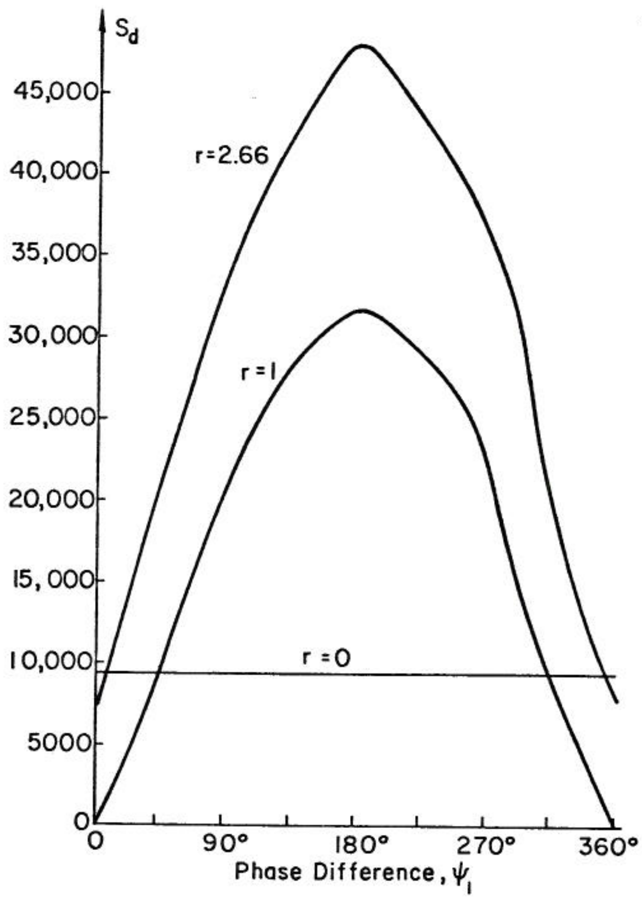

Fig. 3-8. Difference Storage Versus Phase Difrerence ψ_1 Corresponding to $n = 1460$ for Different Amplitude Ratios, r.

$$a\ c\ [n(100)]^{c-1} = b\ d\ [n(100)]^{d-1} \tag{3-25}$$

or

$$(c-d)\ \ln[n(100)] = \ln b + \ln d - \ln a - \ln c\ . \tag{3-26}$$

Case II $\qquad$ p% development

Now $\qquad S_t(p) \approx h\ n^c + S_d(p)$, $\qquad$ (3-27)

while S_s is still given by Eq. (3-22), because it does not depend on the output. Therefore,

$$\frac{dS_t(p)}{dn} \approx h\ c\ n^{c-1} \tag{3-28}$$

Assuming Eqs. (3-24) and (3-28) to be equal at $n = n(p)$, then

$$h\ c\ [n(p)]^{c-1} = b\ d\ [n(p)]^{d-1} \tag{3-29}$$

or

$$(c-d)\ \ln\ [n(p)] = \ln b + \ln d - \ln h - \ln c\ , \tag{3-30}$$

then on dividing Eq. (3-26) by Eq. (3-30)

$$\frac{\ln\ [n(100)]}{\ln\ [n(p)]} = \frac{\ln b + \ln d - \ln a - \ln c}{\ln b + \ln d - \ln h - \ln c} \tag{3-31}$$

Now for the stabilization region of partial regulation to be longer than the full regulation, $n(p)$ should be greater than $n(100)$ i.e., the denominator in Eq. (3-31) should be more than its numerator. Therefore,

$$\ln b + \ln d - \ln h - \ln c > \ln b + \ln d - \ln a - \ln c$$

or

$$-\ln h > -\ln a$$

or

$$a > h \tag{3-32}$$

This means the rate of increase of total storage with n is more for 100% development than for p% development, when p is less than 100% for partial regulation. This holds always true under the above assumptions as per Fig. 3-1. Hence the initial assumption of $n(100) < n(p)$ is also true.

It is thus concluded that the length of stabilization region is inversely proportional to the level of development.

Effect of Dependence of Stochastic Component of Input on Stabilization Region. So far only the effect of changing the rate of increase of total storage with n for different percentages of regulation has been considered on the length of stabilization region. However, since the difference storage depends on the stochastic storage too, so the rate of increase of stochastic storage with n should also be considered for various conditions.

The rate of increase of stochastic storage can be varied by changing the dependence structure of stochastic variable $\varepsilon_{p,\tau}$ of the input process. The difference storage, obtained for the two cases of dependent and independent structure of $\varepsilon_{p,\tau}$, is given in Fig. 3-7. It is found from Table 3-3 and Fig. 3-5 that by removing the dependence from the stationary stochastic component $\varepsilon_{p,\tau}$ of the input process, the rate of increase of the stochastic storage becomes reduced much more than the reduction affected in total storage. Hence for 90% regulation, where the rate of increase of stochastic storage is greater than the rate of increase of total storage for a dependent process, the length of stabilization region should reduce by removing the dependence of the stochastic component of the input process. This is clear in Fig. 3-7, where the difference storage has stabilized at $n = 3650$ against $n = 5000$ for the dependent process.

3-9 Estimation of Difference Storage Capacity

Since difference storage depends on many factors, no general curves can be plotted. However, the good point is that it becomes stable after a small value of n, thereby reducing the length n of daily flow data in order to estimate the total and stochastic storage to compute the difference storage. For instance, for a reservoir with economic life of 100 years, the difference storage capacity could be estimated from a daily flow subseries length of 4 years for 100% development, and 15 years for the 90% development. The following are the two ways by which the value of difference storage capacity can be estimated in practice.

(1) By using the actual daily flow series; and
(2) By simulating the daily flow series.

The method of using the actual data would be applicable only at those sites where the record available is long, so as to give reasonably accurate estimates of the total and stochastic storages.

The method of using the simulated daily flows is applicable for all sites with short record of data. The main advantage is that the procedure does not require the simulation of daily flow series length equal to the economic life of the reservoir. The following are the steps to be followed in estimating the difference storage capacity by this method:

(a) Find the significant harmonics of the periodic mean and periodic standard deviation;
(b) Remove periodicity from the series by $(x_{p,\tau} - \hat{\mu}_\tau)/\hat{\sigma}_\tau$;
(c) Standardize the residual series;
(d) Fit a Markov model to the series obtained, by estimating the autoregressive coefficients from the correlation coefficients of the standardized series;
(e) Simulate the daily flows by Eq. (3-5);
(f) Estimate the total storage by finding the mean maximum deficit of the net input series for a given output;
(g) Estimate the stochastic storage by finding the mean maximum deficit of stochastic net input series;
(h) Difference storage represents the difference of the total and stochastic storages; and
(i) Repeat the above procedure for different values of n, till the value of difference storage becomes approximately constant. This constant value is the difference storage capacity, to be adopted in the design of reservoir with long economic life.

Chapter 4
STOCHASTIC STORAGE

4-1 Definition of Stochastic Storage

Let $x_{p,\tau}$ be the daily discharge at day τ of the year with the earlier given definition of p and τ. It can be represented mathematically by

$$x_{p,\tau} = \mu_\tau + \sigma_\tau \varepsilon_{p,\tau} \qquad (4\text{-}1)$$

In Eq. (4-1), μ_τ is considered as the deterministic input and $\sigma_\tau \varepsilon_{p,\tau}$ as the stochastic input. As regards output, its deterministic portion has already been dealt with in Chapter III; however, its stochastic component if any must be considered. Since output $Y_{p,\tau}$ was assumed to be deterministic in this study, the stochastic output will be assumed zero, which is equal to the expected value of the stochastic input $\sigma_\tau \varepsilon_{p,\tau}$. Hence the stochastic storage could be defined either on the concept of range or on maximum deficit. However, it may be recalled from Chapter III that the total storage has been defined on the concept of maximum deficit and hence the stochastic storage, if defined on any other concept like the range or the adjusted range, would not stabilize the asymptotic value of the difference storage to a constant. To exploit this property of the difference storage in the design of reservoirs, the stochastic storage must be defined on the concept of maximum deficit, which turns out to be more economical than design based on the concept of range even for the case of full regulation. Hurst (1951) states that if the output from a reservoir over a period is kept steady at the mean, the range of the progressive sums of departures from the mean is equivalent to the storage required to keep this constant output. However, it may be claimed that even in the case of a full regulation (alternate expressions are the regulation on the mean or the complete regulation or 100% development) the storage required should be defined on the concept of maximum deficit rather than on the range.

Let consider a realization of the stochastic process as per Fig. 4-1. If the range is accepted as the criterion for design of storage capacities, then the storage required is r_n, i.e., the reservoir empty at b and full at c. This means the operation should start with an initial quantity of water in reservoir equal to V and the future operation would start with the quantity $(V+S_n)$. Since $d_1 < d_n$, there is no point in assuming the reservoir full at c, when one could easily meet the demand through this realization by assuming the reservoir full at a and c_1 and spill the water between the time period of points c_1 and c. The spill is equal to S such that $S = S_n$, and hence the reservoir would start again with the quantity V for the future operation. This assumption of reservoir full at a and c_1 and empty at b is the basis of maximum deficit analysis or the sequent-peak method, which gives the required storage capacity as d_n instead of r_n. In this particular realization $d_n < r_n$, while considering any other realization of this process as per Fig. 4-2, it can be seen that $d_n = r_n$. Thus, it can be said that

$$P[D_n \le R_n] = 1 \qquad (4\text{-}2)$$

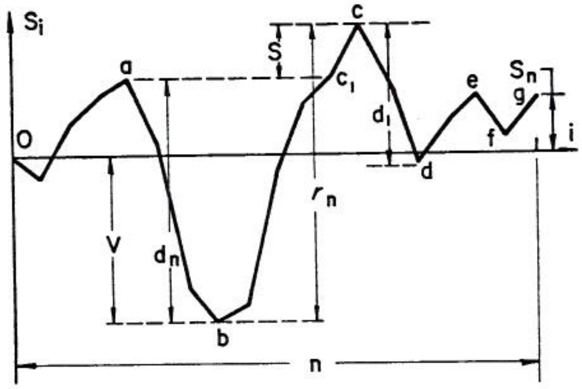

Fig. 4-1. Example Showing the Value of Maximum Deficit Being Less than the Range.

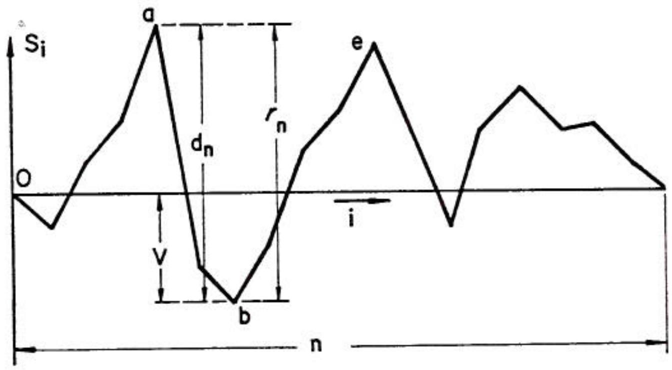

Fig. 4-2. Example Showing the Value of Maximum Deficit Being Equal to the Range.

This proves that the design based on the maximum deficit analysis would require the capacity of a reservoir less or equal to that based on the analysis of the range, and consequently is more economical to apply in practice. Hence in spite of the stochastic output being equal to the expected value of stochastic input, thereby having a case of full regulation, the stochastic storage is defined in this study as the expected maximum deficit of the stochastic net input series.

Since the stochastic net input, $\sigma_\tau \varepsilon_{p,\tau}$, is nonstationary due to the presence in it of a periodic parameter, σ_τ, the general approach to find the stochastic storage would be to find the expected maximum deficit of a stationary process $\varepsilon_{p,\tau}$, if $\varepsilon_{p,\tau}$ is stationary of the second or third order, and then convert the same to produce the stochastic storage for the nonstationary process, $\sigma_\tau \varepsilon_{p,\tau}$.

4-2 Dependence Model of Stationary Stochastic Component of Input

It has been shown by many investigators that the variable $\varepsilon_{p,\tau}$ obtained by removing the periodicity in the mean and the standard deviation is only approximately a second-order stationary dependent or independent time series. General m-th order autoregressive linear dependence models have been used by many investigators (Yevjevich, 1964: Roesner and Yevjevich, 1966; and Quimpo, 1967) for determining the dependence structure of annual, monthly, and daily precipitation and runoff series. Since the daily flow is the input to the reservoir in this study, the stationary stochastic component of daily flow series is assumed to follow a second-order Markov model, as an average case between the use of the first-order model (currently often used in practice), and the likely need for the third and higher order models, namely

$$\varepsilon_{p,\tau} = \alpha_1 \varepsilon_{p,\tau-1} + \alpha_2 \varepsilon_{p,\tau-2} + s\ \eta_{p,\tau} \ , \qquad (4\text{-}3)$$

where s is given by Eq. (3-7), with a_1 and a_2 as the estimates of α_1 and α_2.

The basic question in the model of Eq. (4-3) is whether or not α_1 and α_2 are periodic. Yevjevich (1972) stated that the autoregressive coefficients α_1 and α_2 are usually nonperiodic parameters for the precipitation series and the series of river flows which are mainly produced by rainfall. However, the river flow series of mixed rainfall and snowmelt contribution to runoff may usually have periodic autocorrelation coefficients, $\rho_{k,\tau}$, and consequently periodic autoregressive coefficients $\alpha_{j,\tau}$. Therefore a sensitivity analysis should first be carried out for the expected maximum deficit of Eq. (4-3), by using variation in autocorrelation coefficients as per Fig. 4-3, and then taking their average values to ascertain the effects on the deficit. The autoregressive coefficients are related to autocorrelation coefficients by

$$\alpha_{1,\tau-1} = \frac{[\rho_{1,\tau-1} - \rho_{1,\tau-2}\ \rho_{2,\tau-2}]}{(1 - \rho^2_{1,\tau-2})} \qquad (4\text{-}4)$$

and

$$\alpha_{2,\tau-2} = \frac{[\rho_{2,\tau-2} - \rho_{1,\tau-1}\ \rho_{1,\tau-2}]}{(1 - \rho^2_{1,\tau-2})} \qquad (4\text{-}5)$$

with the ρ-values estimated by the r-values, and by using the r-values in Eqs. (4-4) and (4-5), $a_{1,\tau-1}$ and $a_{2,\tau-2}$ are computed as the estimates of $\alpha_{1,\tau-1}$ and $\alpha_{2,\tau-2}$, respectively.

Two different second-order autoregressive series of model of Eq. (4-3) were generated, one with the periodic autoregressive coefficients given by Eqs. (4-4) and (4-5), and the other with the constant autoregressive coefficients. The expected values of the maximum deficit obtained for the two cases are plotted in Fig. 4-4.

From Fig. 4-4 it can be concluded that the mean of maximum deficit is not affected substantially by the fluctuation of autocorrelation coefficients, whether taken periodic or constant. The study has not been conducted with various shapes of variation of periodic autocorrelation coefficients, because the same has been carried out for the case of expected range by Salas (1972), and he too arrived at the same conclusion.

It is, however, interesting to see as to how the assumption of constant autoregressive coefficients in model of Eq. (4-3) compares with values obtained from the actual data. The study is conducted on the Oconto River. The periodicity of its 40 year daily flow data was removed in the mean and in the standard deviation and the second-order Markov model was fitted to its stochastic component. The details of the analysis are given in Chapter V, in Tables 5-1 through 5-3.

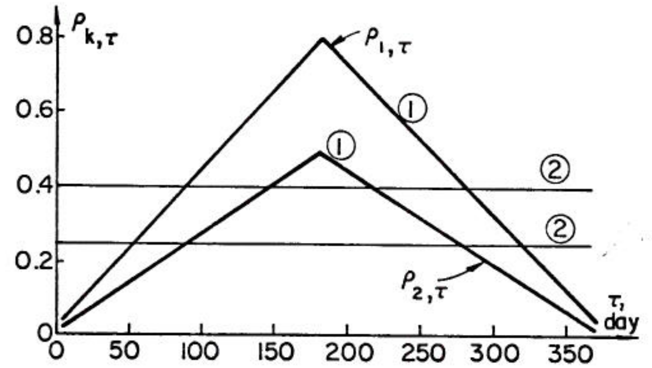

Fig. 4-3. Variation of Correlation Coefficients During the Year: (1) Curves of Periodic Autocorrelation Coefficients, $\rho_{1,\tau} = 0.40$, $s(\rho_{1,\tau}) = 0.23$, $\rho_{2,\tau} = 0.25$, $s(\rho_{2,\tau}) = 0.144$; and (2) Curves of Constant Autocorrelation Coefficients, $\rho_{1,\tau} = 0.40$, and $\rho_{2,\tau} = 0.25$.

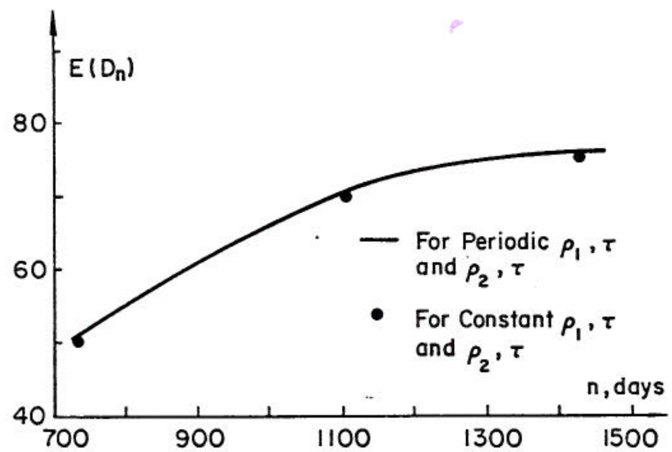

Fig. 4-4 Expected Maximum Deficit for Different Values of n Corresponding to Periodic and Constant Autocorrelation Coefficients.

The second-order stationary series for $n = 730$ with two estimated autoregressive coefficients are simulated and the mean and standard deviation of maximum deficit found as $E(D_n) = 108.41$, and $\sqrt{Var(D_n)} = 56.58$. With the help of the standardized distribution of the maximum deficit given in Fig. 4-8, the value of maximum deficit at 90% confidence level is $D_n(90\%) = E(D_n) + 1.6[Var(D_n)]^{1/2} = 108.41 + 1.6 \times 56.58 = 198.94$.

For the 40-year daily flows of the Oconto River broken into 20 series, each with a 2-year length, the periodicity of each subseries is removed by $(x_{p,\tau} - \hat{\mu}_\tau)/\hat{\sigma}_\tau$ with $\hat{\mu}_\tau$ and $\hat{\sigma}_\tau$ estimated from the whole data. The maximum deficits of the twenty series so obtained are given in Table 4-1. If the actual data comes from the population of Eq. (4-3), with the constant autoregressive coefficients, then the 90% of the subseries of actual data should have the maximum deficit less than the value of 198.94, obtained above. From Table 4-1 it is found that only three values are greater than 198.94 with the third value of 200.34 close to it. Hence it can be assumed that at 90% confidence level the actual data comes of the population of Eq. (4-3). It is then concluded that the further study could be carried out only with the constant autocorrelation coefficients, and consequently autoregressive coefficients.

Table 4-1. Maximum Deficit of Different Series of Length 730 Days of Stochastic Component of Actual Data of Oconto River

Series No.	Maximum Deficit
1	150.99
2	131.52
3	140.92
4	166.72
5	238.82
6	196.74
7	160.49
8	137.71
9	174.39
10	145.93
11	113.48
12	196.87
13	177.55
14	153.97
15	143.68
16	190.22
17	101.82
18	200.34
19	178.54
20	224.52

4-3 Determination of Stochastic Storage

Salas (1972) has found that the stochastic storage, defined as the expected range of the stochastic net input, is a function of mean and standard deviation of σ_τ, the dependence ρ of the stochastic component of input and n. Therefore the stochastic storage can in principal be expressed by

$$E(D_n) \simeq f[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \quad , \tag{4-6}$$

where $\bar{\sigma}_\tau$ and $s(\sigma_\tau)$ denote the mean and the standard deviation of the periodic standard deviation, while ρ denotes the dependence of the stationary stochastic component $\varepsilon_{p,\tau}$. The dependence expressed by ρ may be of any Markovian type model. A function $f_1(1,0,0)$, similar to that of Eq. (4-6), may be defined with the expected maximum deficit of this process having $\bar{\sigma}_\tau$ = 1, $s(\sigma_\tau) = 0$ and $\rho = 0$, i.e., the expected maximum deficit of an independent process with the standard deviation of 1.

To obtain the stochastic storage of a nonstationary process given by Eq. (4-6), the basic hypothesis is to separate the effect of dependence in $\varepsilon_{p,\tau}$ from the effect of nonstationarity due to the periodic σ_τ. With this objective in mind, the expected maximum deficit may be investigated for the following four types of functions,

$$f_1 = f_1(1,0,0) \quad ,$$

$$f_2 = f_2(1,0,\rho) \quad ,$$

$$f_3 = f_3[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ 0] \quad ,$$

and

$$f_4 = f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \quad , \tag{4-7}$$

where f_1, f_2, f_3, and f_4 can be interpreted in relation to Eq. (4-6) similarly as the interpretation of $f_1(1,0,0)$ given above. Then the stochastic storage in the form of the function f_4 can be expressed mathematically in accordance with the above hypothesis by

$$f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \simeq \bar{\sigma}_\tau[f_2(1,0,\rho) - f_1(1,0,0)] + $$
$$+ f_3[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ 0] \tag{4-8}$$

It may be noted that the expression $[f_2(1,0,\rho) - f_1(1,0,0)]$ of Eq. (4-8) represents the increase in the expected maximum deficit due to dependence in the stationary stochastic component $\varepsilon_{p,\tau}$, with a standard deviation equal to one, while $f_3[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ 0]$ represents the effect of nonstationarity by way of the expected maximum deficit of a nonstationary independent process.

From Eq. (4-8) it is clear that if the values for $f_2(1,0,\rho)$ and $f_3[\bar{\sigma}_\tau, s(\sigma_\tau), 0]$ are known, then the stochastic storage in the form of $f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho]$ can be estimated. The first step is to find the value of $f_2(1,0,\rho)$, which is the expected maximum deficit of a stationary dependent process with mean zero and standard deviation unity for the case of full regulation.

Expected Maximum Deficit of Markov Models. Francisco L.S. Gomide (1975) dealt with the theoretical analysis of maximum deficit for the independent process, and gave the asymptotic expected maximum deficit as

$$E(D_n) = 1.2533\sqrt{n} \quad , \tag{4-9}$$

and

$$Var(D_n) = 0.2611\ n \tag{4-10}$$

He stated that the next step, after solving the case of the independent process, was to extend the theory of Markov chains to the case of seasonal and correlated inputs. The extension to the case of seasonality in input was merely mentioned, and the extension to the correlated inputs was made only for the very simple cases because of the limitation found by Lloyd, namely, of a drastic increase in the size of matrices involved. Thus for a dependent process, the expression for the expected maximum deficit is not available in closed form, even for the simple first-order model, and less so for the higher-order Markov models, which the stationary stochastic component of the daily flow series may follow.

Thus, the only alternative left was to go through the generation procedure. Before attempting any generation, the first question to consider was whether both cases are of interest, the finite n and the asymptotic case. Whenever the design of a reservoir capacity is made for within-the-year fluctuations, the problems with either monthly, weekly or daily periodicity in inputs and outputs must be considered. Even in monthly periodicity in input, the value of n is 1200 for a 100-year reservoir economic life, for which the asymptotic results should be well applicable. Since this study is concerned with the design of a large reservoir with daily flow data, the asymptotic results are assumed well applicable.

Representing the m-th order Markov model by

$$z_t = \sum_{j=1}^{m} a_j z_{t-j} + \eta_t , \qquad (4\text{-}11)$$

the samples of Markov models of various orders are generated by using Eq. (4-11), with the autoregressive coefficients so chosen that they satisfy the conditions of stationary processes, namely for second-order Markov models

$$(a_1 + a_2) < 1$$

$$(a_1 - a_2) > -1$$

$$-1 < a_2 < 1 \qquad (4\text{-}12)$$

Assuming the 50-year economic life of reservoirs, the value of n for daily flow series would be approximately 18,000. Hence the thirteen series are generated for three different orders of Markov linear models, all three having the same sum of the autoregressive coefficients. The distribution of η_t is assumed normal (0,1). The distribution of η_t is in fact not important as per Feller's result of asymptotic distribution of the range being independent of the distribution of the underlying process, because the partial sums of independent random variables with finite variance are asymptotically normally distributed. The results obtained from generated samples are given in Tables 4-2 and 4-3.

It can be inferred from Tables 4-2 and 4-3, that the sum rather than the individual values of autoregressive coefficients determine the distribution of the maximum deficit and the range. The results for the third-order model are given in Table 4-4. This latter table also supports the conclusion.

The other interesting result is that the values of the range and the maximum deficit, etc., for each series generated with the same random numbers but with the two different autoregressive schemes of the same sums of autoregressive coefficients, are almost the same (refer to Table 4-5).

Example of Application on the Tioga River. The daily flow data of the Tioga River are taken to further study the hypothesis that the $\sum a_j$ is a determining factor and not its individual values. Harmonic analysis is performed on the mean and standard deviation of 365 daily values of the year, following the procedure given by Yevjevich (1972a). The results are presented in Tables 4-6 and 4-7. After removing the periodicity in the daily flows, its stochastic component is fitted by the second-order Markov model. The

Table 4-2. Percentile of Distribution of Maximum Deficit for Different Sums of the Autoregressive Coefficients. Markov Model Used is $z_t = a_1 z_{t-1} + a_2 z_{t-2} + \eta_t$ and n = 18,000

	0.5			0.7			0.9		
S.No.	a_1	a_2	Percentile of Distribution of Maximum Deficit	a_1	a_2	Percentile of Distribution of Maximum Deficit	a_1	a_2	Percentile of Distribution of Maximum Deficit
1	0.4	0.1	448.29	0.35	0.35	741.88	0.5	0.4	2163.08
2	0.3	0.2	448.02	0.80	- 0.10	746.50	0.6	0.3	2169.64
3	0.5	0.0	448.53	-	-	-	-	-	-

Table 4-3. Percentile of Distribution of Surplus, Deficit and Range of Different Markov Models with the Same Sum of Autoregressive Coefficients. Markov Model Used is $z_t = a_1 z_{t-1} + a_2 z_{t-2} + \eta_t$.

Autoregressive Coefficients a_1	a_2	Order of Markov Model	Percentile of Distribution of Surplus	Percentile of Distribution of Deficit	Percentile of Distribution of Range
0.4	0.1	2 nd	319.992	-413.3778	510.8439
0.5	0.0	1 st	320.384	-412.9498	511.1853

Table 4-4. Percentile of Distribution of Maximum Deficit of Third-Order and Second-Order Markov Model with Same Sum Equal to 0.7, of Autoregressive Coefficients. Markov Model Used was $z_t = a_1 z_{t-1} + a_2 z_{t-2} + \eta_t$ and n = 18,000

S. No.	a_1	a_2	a_3	Percentile of Distribution of Maximum Deficit
1	0.35	0.35	-	741.88
2	0.80	(-) 0.10	-	746.50
3	0.30	0.20	0.20	742.09

results are given in Table 4-8. Two schemes are used, the second-order model with a_1 and a_2 of Table 4-8, and the other with the first-order model, with $r_1 = a_1 + a_2$. The obtained distributions of range and maximum deficit for the two schemes are shown in Figs. 4-5 and 4-6 respectively. The mean range and the mean maximum deficit for the two cases are given in Tables 4-9 and 4-10 respectively.

It can be concluded from Tables 4-9 and 4-10 that the parameters of distributions, even for the distributions of the range and the maximum deficit, can be obtained by equating the higher-order Markov model to its equivalent first-order model with $\rho_1 = \alpha_1 + \alpha_2 + \ldots + \alpha_m$, so that $|\rho_1| \leq 1$. This constraint on ρ_1 should be considered in detail before applying this approach. Considering a simple case of a second-order model, the conditions for the stationarity of the

Table 4-6. Results of Harmonic Analysis for Daily Flow of the Tioga River

Harmonic		Values of Coefficients of Harmonics	
	A	-1274.70	-1286.60
	B	245.30	485.00
1	Explained Variance	0.5933	0.3574
	A	613.50	565.10
	B	268.60	287.50
2	Explained Variance	0.1579	0.0760
	A	- 501.40	- 386.30
	B	- 42.10	168.10
3	Explained Variance	0.0891	0.0336
	A	238.80	168.50
	B	- 194.40	- 267.00
4	Explained Variance	0.0334	0.0188
	A	- 219.60	- 243.60
	B	129.90	320.30
5	Explained Variance	0.0229	0.0306
	A	89.80	117.50
	B	- 136.00	- 207.20
6	Explained Variance	0.0094	0.0107
Variance Explained by Six Harmonics		0.9060	0.5271

Table 4-5. Comparison of Surplus, Deficit, Range, and Maximum Deficit of Different Series of n = 18,000 with Same Independent Standard Normal Random Numbers, But with Different Order of Markov Models with the Same Sum of Autoregressive Coefficients.

Series No.	Surplus		Deficit		Range		Maximum Deficit	
	2nd Order	1st Order	2nd Order	1st Order	2nd Order	1st Order	2nd Order	1st Order
1	226.90	226.83	-175.22	-175.61	402.13	402.44	257.70	258.23
2	49.54	49.70	-235.44	-235.76	284.98	285.45	251.06	251.57
3	41.48	41.99	-413.34	-412.89	454.82	454.87	454.82	454.87
4	124.78	124.89	-178.83	-178.90	303.62	303.79	292.84	293.23
5	318.38	318.76	- 12.86	- 12.76	331.25	331.53	268.67	268.92
6	90.63	90.26	-321.02	-321.73	411.65	412.00	361.98	362.98
7	56.11	55.71	-182.14	-183.10	238.25	238.81	238.25	238.81
8	256.96	257.26	- 58.61	- 58.92	315.56	316.17	247.90	248.26
9	153.37	153.21	-259.54	-259.92	412.91	413.13	412.91	413.13
10	239.66	241.04	- 22.06	- 21.92	261.72	262.95	213.03	212.83
11	374.74	375.56	-189.90	-189.81	564.64	565.37	231.29	231.30
12	143.15	143.64	-366.11	-365.96	509.26	509.59	366.11	365.90
13	23.84	24.02	-414.66	-414.98	438.50	439.00	438.50	439.00

Note : (1) In second-order Markov model $a_1 = 0.40$, and $a_2 = 0.10$; and

(2) In first-order Markov model $a_1 = 0.50$, and $a_2 = 0.00$.

Table 4-7. Number of Harmonics Used to Represent μ_τ and σ_τ of Tioga River

μ_τ			σ_τ		
Number of Harmonics	Percent of Explained Variance	Total Explained Variance due to Six Harmonics	Number of Harmonics	Percent of Explained Variance	Total Explained Variance due to Six Harmonics
3	0.8403	0.9060	6	0.5271	0.5271

Table 4-8. Fitting the Second-order Autoregressive Model to Standardized Stochastic Component of Daily Flow Data of Tioga River

a_1	a_2	Variance of Residual Series	
		Theoretical	Computed
0.6093	- 0.02967	0.6493	0.6486

Table 4-9. Comparison of Mean Range of Second-order and Equivalent First-order Markov Model of the Tioga River Daily Flows for n = 1440

Model						
Second-order				First-order		
a_1	a_2	Mean Range	Standard Deviation of Range	r_1	Mean Range	Standard Deviation of Range
0.6093	- 0.02967	198.09	58.30	0.5796	197.95	58.29

Table 4-10. Comparison of Mean Maximum Deficit of Second-order and Equivalent First-order Markov Model of the Tioga River Daily Flows for n = 1440

Model						
Second-order				First-order		
a_1	a_2	Mean Maximum Deficit	Standard Deviation of Maximum Deficit	r_1	Mean Maximum Deficit	Standard Deviation of Maximum Deficit
0.6093	- 0.02967	106.58	-	0.5796	106.44	-

process are that the values of a_1 and a_2 should lie in the triangle ABC, Fig. 4-7. Let us assume a second-order model corresponding to the point B, where a_1 = -2 and a_2 = -1. Its hypothetical equivalent first-order Markov model should have $r_1 = -2 - 1 = -3$. Since $|r_1|$ cannot be more than 1, hence this model has no meaning. It is thus concluded that this equivalent conversion is applicable to the second-order models, whose autoregressive coefficients lie in the region ADEC. Fortunately, this restriction of applicability to the region ADEC does not limit the application of this approach to practical problems due to the following reason.

In storage problems the input is always a runoff series, which may be monthly, weekly or daily flows. Since hydrologic processes are persistent, the first serial correlation should at least be always positive, and hence a_1 for $r_1 = a_1/(1-a_2)$ should also be always positive. This indicates that in practical application, the values of a_1 and a_2 will not lie in an area ABE of Fig. 4-7. Thus, the conclusion of converting the higher-order Markov model to its equivalent first-order model is applicable to all storage problems of practical interest.

Having derived this conclusion of converting the higher-order Markov models to its equivalent first-

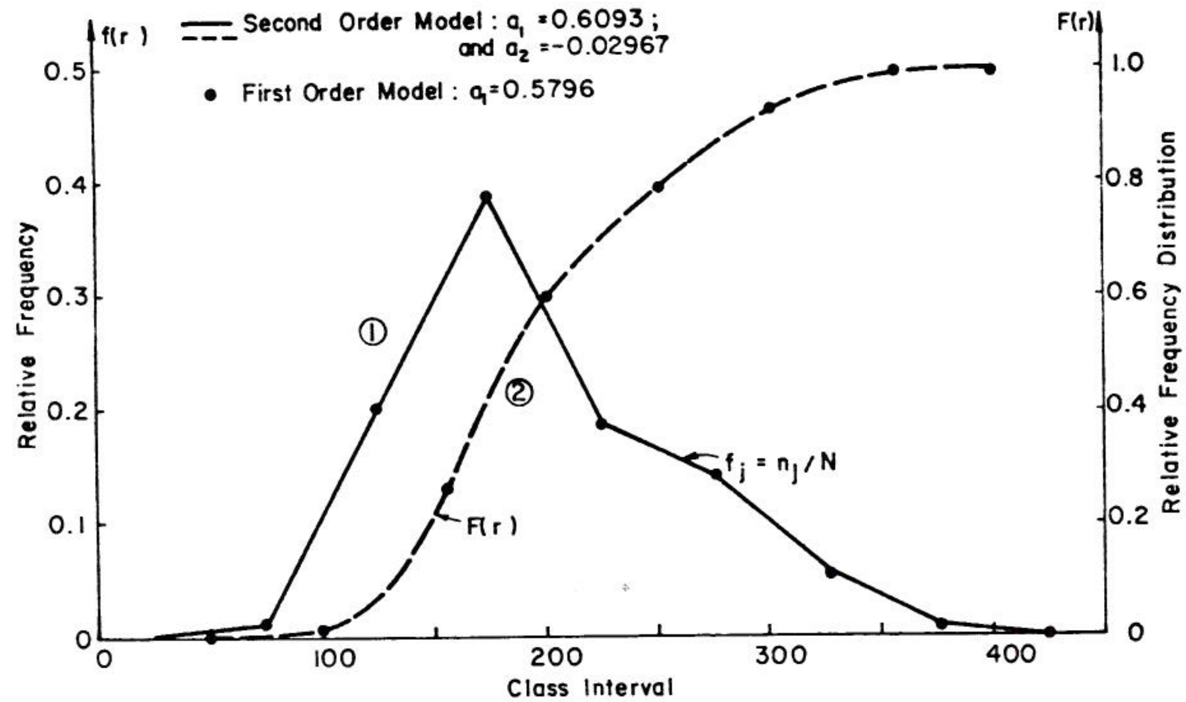

Fig. 4-5. Polygon of Relative Frequency (1) and Relative Frequency Distribution (2) of Range for Second-Order and Equivalent First-Order Markov Model of Tioga River for n = 1440.

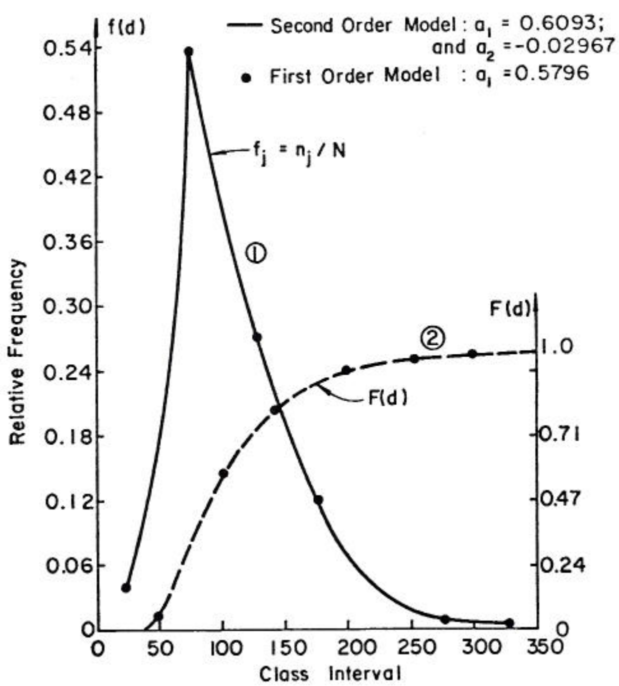

Fig. 4-6. Polygon of Relative Frequency (1), and Relative Frequency Distribution (2) of Maximum Deficit for Second-Order and Equivalent First-Order Markov Model of Tioga River for n = 1440.

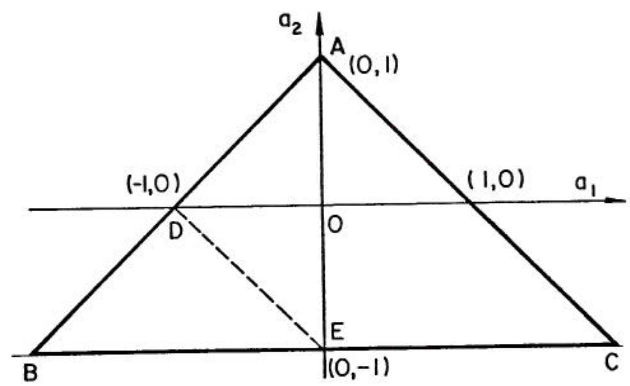

Fig. 4-7. Region for a_1 and a_2 for the Second-Order Stationary Markov Model.

order models for the determination of distributions of different statistics of partial sums through the sample generation approach, the next step was to prove this result by a theoretical analysis.

Theoretical Analysis. Representing the second-order Markov model by

$$z_t = a_1 z_{t-1} + a_2 z_{t-2} + \eta_t \tag{4-13}$$

or

$$\eta_t = z_t - a_1 z_{t-1} - a_2 z_{t-2} \quad . \tag{4-14}$$

For

$$\eta_1 = z_1, \ \eta_2 = z_2, \text{ and } \eta_3 = z_3 - a_1 z_2 - a_2 z_1 \quad ,$$

and

$$\eta_4 = z_4 - a_1 z_3 - a_2 z_2, \ldots, \eta_i = z_i - a_1 z_{i-1} - a_2 z_{i-2} ,$$

let

$$S_{i,z} = \sum_{j=1}^{i} z_j , \text{ and } S_{i,\eta} = \sum_{j=1}^{i} \eta_j .$$

Then summing up η_i's and their right sides, then

$$S_{i,\eta} = z_i + (1-a_1)z_{i-1} + (1-a_1-a_2) \sum_{j=1}^{i-2} z_j + (1-a_2)z_1 . \tag{4-15}$$

For i very large, the end effects in Eq. (4-15) being too small can be neglected and it can be written as

$$S_{i,\eta} \approx (1-a_1-a_2)S_{i,z}$$

or

$$S_{i,z} \approx \frac{S_{i,\eta}}{(1-a_1-a_2)} \tag{4-16}$$

If initially the m-th order Markov model is assumed in Eq. (4-13), namely

$$z_t = \sum_{j=1}^{m} a_j z_{t-j} + \eta_t , \tag{4-17}$$

then Eq. (4-16) becomes

$$S_{i,z} \approx \frac{S_{i,\eta}}{(1 - \sum_{j=1}^{m} a_j)} \tag{4-18}$$

Since $S_{i,\eta}$ or $S_{i,z}$ represent a new stochastic variable, the cumulative sum of deviations of η or z, respectively, it is a function of i and can be represented as $S_i = f(i)$ of a discrete series of the continuous variable. It is this stochastic variable which determines the properties of the surplus, deficit, range, and maximum deficit.

It is evident from Eq. (4-18) that all properties of S_i, like the range and maximum deficit of the m-th order Markov model, are inversely proportional to $(1 - \sum_{j=1}^{m} a_j)$. Considering the second-order model, it is obvious from Eq. (4-16) that the greater the sum of $(a_1 + a_2)$, the greater would be the value of the range and the maximum deficit, which is intuitive also. It may be concluded also from the above result that the distributions of the range and the maximum deficit for large i are functions of the sum $(a_1 + a_2)$ rather than the individual values of a_1 and a_2, thereby supporting the earlier conclusion obtained through the sample generation approach.

If the initial model in Eq. (4-13) is taken to be the first-order Markov model

$$z_t = \rho_1 z_{t-1} + \eta_t , \tag{4-19}$$

then the asymptotic expected maximum deficit of this model is given by Eq. (4-18) as the expected maximum deficit of the independent process with the mean zero and standard deviation unity, divided by $(1 - \rho_1)$. Hence,

$$E(D_n) = \frac{1.2533 \sqrt{n}}{1 - \rho_1} \tag{4-20}$$

The variance z of Eq. (4-19) is $1/(1 - \rho_1^2)$. Hence, the expected maximum deficit of model represented by Eq. (4-19) with unit variance is obtained by dividing Eq. (4-20) by $[1/(1 - \rho_1^2)]^{1/2}$. Then

$$E(D_n) = \frac{1.2533 \sqrt{n} \, [1 - \rho_1^2]^{1/2}}{(1 - \rho_1)}$$

$$E(D_n) = \frac{1.2533 \sqrt{n(1 + \rho_1)}}{\sqrt{(1 - \rho_1)}} \tag{4-21}$$

It may be noticed that the factor $[(1 + \rho_1)/(1 - \rho_1)]^{1/2}$ is the same as for the case of range and adjusted range. This should be so, as all these statistics relate to the same stochastic variable, the cumulative partial sums.

It may also be noted from Eq. (4-18) that if $S_{i,\eta}$ is the same, i.e., the same generated random numbers of same length are used for two different autoregressive processes of Eq. (4-17), but in a way that the sum of the autoregressive coefficients is the same for both series, then the values of the range and the maximum deficit should be the same for both series, thereby supporting the results obtained in Table 4-5.

If Eq. (4-13) is modified to be

$$z_t = a_1 z_{t-1} + a_2 z_{t-2} + s \, \eta_t , \tag{4-22}$$

then Eq. (4-18) becomes

$$S_{i,z} = \frac{s \, S_{i,\eta}}{(1 - \sum_{j=1}^{m} a_j)} , \tag{4-23}$$

indicating that the asymptotic expected value and the variance of any statistic of partial sums for the Markov dependent models are functions of the respective asymptotic values of an independent process. Thus for the m-th order Markov model the asymptotic expected range is

$$E(R_n) = 1.5958 \, \beta \sqrt{n} , \tag{4-24}$$

$$\mathrm{Var}(R_n) = \beta^2 \, 4n(\ln 2 - \tfrac{2}{\pi}) , \tag{4-25}$$

where

$$\beta = \frac{s}{(1 - \sum_{j=1}^{m} a_j)} . \tag{4-26}$$

Thus the parameters of the distribution of any statistic of partial sums for the Markov dependent models can be obtained directly from their respective asymptotic values of the independent process.

The above analysis leads to the following important conclusions:

(1) The distribution of all the statistics of partial sums for the dependent and independent process are determined from the stochastic variables $S_{i,z}$

and $S_{i,\eta}$ respectively. It is, therefore, clear that the asymptotic distributions of these statistics for the dependent Markov models should be functions of the asymptotic distribution of the respective statistics of the independent process. This explains why Gomide (1975) found that the standardized distribution of the range and the standardized distribution of the maximum deficit of Markovian inputs tends to the standardized asymptotic distribution of the independent process. This conclusion emphasizes the importance of previous studies in the range analysis devoted to the approximate expressions of the first two moments of R_n for the correlated inputs (Yevjevich, 1967);

(2) The analysis presented herein lays the emphasis on the study of only the first-order Markov models because all the higher-order Markov models can be reduced to their equivalent first-order model for the analysis of storage problems of practical interest, where one is mainly concerned with the asymptotic distributions of different statistics of partial sums; and

(3) Procedure developed for the calculation of the expected maximum deficit is applicable to all rivers following the Markovian models.

Having derived the theoretical expression for the calculation of $f_2(1,0,\rho)$, the expected maximum deficit of a stationary dependent process, either by *converting* the higher-order Markov model to its equivalent first-order Markov model or by getting it *directly* through the expression of asymptotic expected maximum deficit for the i.i.d. process by means of a coefficient β given by Eq. (4-26), Eq. (4-8) can now be modified. If the stochastic component of the daily flows follows the m-th order Markov model,

$$\varepsilon_{p,\tau} = \sum_{j=1}^{m} \alpha_j \varepsilon_{p,\tau-j} + \eta_{p,\tau} \quad , \tag{4-27}$$

then its first-order equivalent model is

$$\varepsilon_{p,\tau} = \rho_1 \varepsilon_{p,\tau-1} + \eta_{p,\tau} \quad ,$$

with

$$\rho_1 = \sum_{j=1}^{m} \alpha_j \quad ,$$

and

$\sigma_{\varepsilon,m}$ = the standard deviation of the m-th order Markov model.

Then

$$f_2(1,0,\rho) = \frac{f_1(1,0,0)(1+\rho_1)^{1/2}\,\sigma_{\varepsilon,1}}{(1-\rho_1)^{1/2}\,\sigma_{\varepsilon,m}} \tag{4-28}$$

Equation (4-8) becomes

$$f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \simeq \bar{\sigma}_\tau\left[\frac{f_1(1,0,0)(1+\rho_1)^{1/2}\,\sigma_{\varepsilon,1}}{(1-\rho_1)^{1/2}\,\sigma_{\varepsilon,m}} - f_1(1,0,0)\right] + f_3(\bar{\sigma}_\tau,\ s(\sigma_\tau),\ 0) \quad . \tag{4-29}$$

Now to find the value of $f_4[\bar{\sigma}_\tau, s(\sigma_\tau), \rho]$, the only unknown left in Eq. (4-29) is $f_3[\bar{\sigma}_\tau, s(\sigma_\tau), 0]$, which is the expected maximum deficit of a nonstationary but independent process. Before getting its value, let the expression be first developed for the expected range of a nonstationary i.i.d process.

Expected Range of Nonstationary Independent Process. Yevjevich (1972) has given by conjecture that the expected range of the independent normal process of Eq. (2-6) may be rewritten in the form

$$E(R_n) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} [(\mathrm{Var}\ S_i)^{1/2}]/i \tag{4-30}$$

where Var S_i is the variance of partial sums. Let consider a nonstationary independent net input process with periodic standard deviation σ_τ such that $\tau = 1,2,\ldots,\omega$ and $i = p\omega$, with the usual meaning of p and ω. Then

$$S_i = \sigma_1\eta_1 + \sigma_2\eta_2 + \cdots + \sigma_\omega\eta_\omega + \sigma_1\eta_{\omega+1} + \cdots + \sigma_\omega\eta_{2\omega} + \sigma_1\eta_{(p-1)\omega+1} + \cdots + \sigma_\omega\eta_{p\omega} \quad .$$

Therefore,

$$\mathrm{Var}\ S_i = \sigma_1^2\ \mathrm{Var}\ \eta_1 + \sigma_2^2\ \mathrm{Var}\ \eta_2 + \cdots + \sigma_\omega^2\ \mathrm{Var}\ \eta_\omega + \sigma_1^2\ \mathrm{Var}\ \eta_{\omega+1} + \cdots + \sigma_\omega^2\ \mathrm{Var}\ \eta_{p\omega} \quad .$$

Since η_1 and η_2 etc. are all i.i.d. with the same mean zero and variance unity, hence

$$\mathrm{Var}\ S_i = p[\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_\omega^2]\ \mathrm{Var}\ \eta \tag{4-31}$$

Let

$$\sigma_n^2 = \left[\frac{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_\omega^2}{\omega}\right] \tag{4-32}$$

Then Eq. (4-31) can be rewritten as

$$\mathrm{Var}\ S_i = p\omega\ \sigma_n^2\ \mathrm{Var}\ \eta = i\ \sigma_n^2 \quad , \tag{4-33}$$

as Var $\eta = 1$, and $p\omega = i$. Substituting Var S_i from Eq. (4-33) in Eq. (4-30)

$$E(R_n) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \frac{i^{1/2}\,\sigma_n}{i}$$

$$E(R_n) = \sqrt{\frac{2}{\pi}}\,\sigma_n \sum_{i=1}^{n} i^{-1/2} \tag{4-34}$$

This gives that the expected range of a nonstationary independent process can be obtained by multiplying the expected range of stationary independent process by an equivalent standard deviation σ_n given by Eq. (4-32). Analogously applying it to the case of maximum deficit it may be written

$$f_3[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ 0] = \sigma_n\ f_1(1,0,0) \tag{4-35}$$

Unlike the case of the range, the effect of increasing the variance of the net input process with the same expectation is not necessarily to increase the expected maximum deficit, hence the conjecture of modifying Eq. (2-6) to Eq. (4-30) is not exactly true in this case. But since σ_τ in the net input process of $\sigma_\tau \eta_\tau$ does not change the expectation from zero, it is still a case of full regulation, for which the expected maximum deficit is known to increase with $\sqrt{n}$. Hence $f_1(1,0,0)$ is taken as $C\sqrt{n}$, where C is a constant, and Eq. (4-29) is modified as

$$f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \simeq C\sqrt{n}\left[\frac{\bar{\sigma}_\tau (1+\rho_1)^{1/2}\ \sigma_{\varepsilon,1}}{(1-\rho_1)^{1/2}\ \sigma_{\varepsilon,m}} - \bar{\sigma}_\tau + \sigma_n\right] \quad (4\text{-}36)$$

It is, therefore, proposed to estimate the value of C by an actual sample generation for a few n values in the range of 5000-7000 for daily flows, and use the average value of C for the determination of stochastic storage for any n corresponding to the economic life of a reservoir, so as to obtain the least error in the calculated stochastic storage. The determination of the average value of C does not need any extra generation effort as the same is required to find out the difference storage capacity as shown in Chapter III.

The method developed to find the first two moments of the distribution of any statistic of partial sums for the higher-order Markov models, can now be used to find the asymptotic distribution of the maximum deficit.

4-4 Asymptotic Distribution of the Maximum Deficit for Higher-Order Markov Model

As described in Chapter I, the exact distributions for large n of *standardized* maximum deficit of partial sums of Markovian net inputs tend to the asymptotic distribution of the *standardized* maximum deficit of the independent process. The standardized asymptotic distribution of the independent process, as given by Gomide (1975), is shown in Fig. 4-8. Thus Fig. 4-8 when corrected for the first two moments, can give the asymptotic distribution of maximum deficit for the Markovian type inputs.

Let apply it to the case of the Tioga River. The dependence of the stationary stochastic component of daily flows has been approximated by the second-order Markov model. The values of autoregressive coefficients as given in Table 4-8 are $a_1 = 0.6093$ and $a_2 =$ (-) 0.02967, with the model

$$z_t = 0.6093\ z_{t-1} - 0.02967\ z_{t-2} + \eta_t \quad . \quad (4\text{-}37)$$

Let $\sigma_{z,m}$ denote the standard deviation of m-th order Markov model, then

$$\sigma_{z,m} = \frac{\text{Var}\ \eta}{\left[1 - \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i\ \alpha_j\ \rho|i-j|\right]^{1/2}} \ , \quad (4\text{-}38)$$

with the α's estimated by a's and ρ values estimated by r values. Thus,

$$\sigma_{z,2} = [1/0.64927]^{1/2} = 1.24 \quad . \quad (4\text{-}39)$$

The distribution of the maximum deficit of the net input process of Eq. (4-37) can be obtained by the equivalent first-order Markov model, with $r_1 = 0.6093 - 0.02967 = 0.5796$, or

$$z_t = 0.5796\ z_{t-1} + \eta_t \ , \quad (4\text{-}40)$$

with $\sigma_{z,1} = [1/(1-0.5796^2)]^{1/2} = 2.0469$.

The asymptotic expected value and the standard deviation of the maximum deficit of series given by Eq. (4-37) is

$$E(D_n) = 1.2533\ \sqrt{n}\ \sigma_{z,1}\ [(1+\rho_1)/(1-\rho_1)]^{1/2} = 266.0641$$

with

$$\sqrt{\text{Var}(D_n)} = \sqrt{0.2611\ n(\sigma_{z,1})^2\ (1+\rho_1)/(1-\rho_1)} = 108.8024$$

Thus the above computed asymptotic expected value and the standard deviation of the maximum deficit for the first-order Markov model of Eq. (4-40) are also the respective values for the second-order model of Eq. (4-37). Then the distribution of the maximum deficit of the series of Eq. (4-37) can be obtained as Fig. 4-9

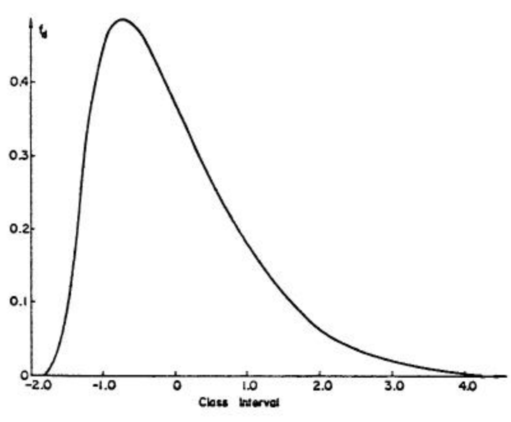

Fig. 4-8. Asymptotic Distribution of $[D_n - E(D_n)]$ $\sqrt{\text{Var}\ D_n}$ for i.i.d. Process, [after Gomide (1975, Figure 6.6)].

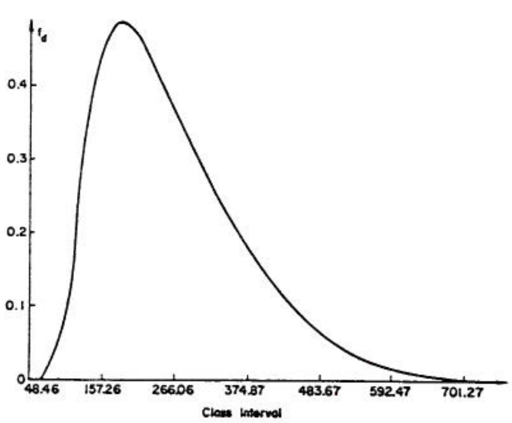

Fig. 4-9. Distribution of Maximum Deficit of Second-Order Markov Model for the Tioga River Daily Flows, with $a_1 = 0.6093$ and $a_2 =$ (-) 0.02967.

by converting the base of Fig. 4-8. This procedure is applicable to any statistic of the cumulative partial sums, such as the range, surplus, the deficit, etc.

4.5 Procedure for Determining the Stochastic Storage

The following are the steps in estimating the stochastic storage for a reservoir to be designed with a long economic life:

(a) Remove the periodicity in daily flow data by means of harmonic analysis, (Yevjevich, 1972a);

(b) Obtain the standardized stochastic stationary series and determine the mean, $\bar{y}$, and standard deviation, s_y;

(c) Fit the suitable Markov linear model to the series obtained under step (b);

(d) The river flows can be generated by using Eq. (3-5), and stochastic storage calculated for a few values of n in the range of 5000-7000 for daily flows;

(e) The theoretical value of stochastic storage for a particular n is obtained by multiplying the value of Eq. (4-36) by s_y;

(f) The theoretical value of stochastic storage obtained under step (e) may be equated to the value obtained under step (d) and the values of C are calculated for all n values. All values of C so obtained will be nearly constant, so that their average value may be adopted for further calculations; and

(g) The stochastic storage for a reservoir can now be calculated by Eq. (4-36) with the help of an average value of C calculated under step (f).

Chapter 5
APPLICATION OF THE DEVELOPED METHOD

The method developed for designing reservoirs with daily flow data is applied to a river, to show how it would work. The 40 years of daily data of the Oconto River near Gillett, Wisconsin are used as an example of application.

5-1 Source of Data and Evaluation of Their Reliability

The U.S. Geological Survey is responsible to gather and publish stream flow data of daily flows for most of the rivers in the United States. During the early observations, the mean daily flows were obtained from daily mean gauge readings on staff gauges. The early records were, therefore, affected by the frequency of observations during a day. The advent of continuous water stage recorders resulted in the replacement of staff gauges. The flows were then calculated by converting the daily mean gauge heights by means of stage-discharge rating curves.

In case the stage-discharge rating curve is subject to change due to frequent or continuous alterations in the physical features of the control, the mean daily discharge is determined by the shifting control method, which involves the application of correction factors, based on individual measurements. This method is also used to correct for temporary changes in the control section due to debris or aquatic growth.

During early stages the crudeness of instrumentation was further aggravated by the lack of sufficient personnel to make frequent observations. This thus required, in some instances, the estimation of unmeasured flows by correlation procedure before publishing the actual data. The perennial problem of ice reducing the area of the control section during winters was another source of error.

The records published by U.S. Geological Survey are classified as excellent, good, fair or poor, depending on whether the errors in them are less than 5, 10, or 15% or greater than 15% respectively.

5-2 Criteria in Selection of Case Station for the Test Study

The basic criteria in selecting a gauge station is to obtain homogeneous records. Gauge stations influenced by significant alterations in the form of diversions or flow regulations upstream through irrigations, diversions or construction of reservoirs, are automatically excluded. Minor diversions up to the maximum of 1% of the average annual flow is tolerated. Station has in fact to be selected on the basis of its virginity of flow. The absence of short term trends is also postulated, in spite of the fact that extensive agricultural use of land, among other things, can cause perceptible trends in runoff.

5-3 The Selected Station

With the above restrictions in mind, the daily flow record available on tape for 19 rivers of the United States was examined. The station number 4.0710 with the drainage area 678 sq miles, of the Oconto River near Gillett, Wisconsin was found to have the accuracy of record classified as good.

The graph of mean daily flows in cubic feet per sec (cfs), obtained by taking the average flow for each day of the year over the 40 years of record for the station, is given in Fig. 5-1. Similarly the standard deviations about the mean daily values are also plotted in Fig. 5-1.

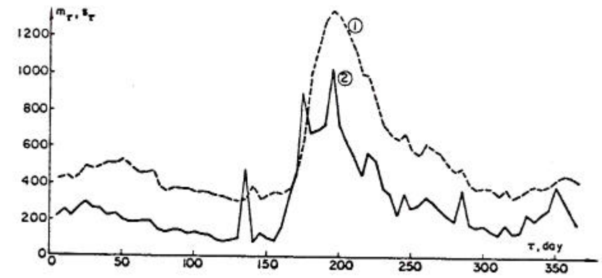

Fig. 5-1. Mean Daily Flows in cfs of Annual Hydrograph (1), and Daily Standard Deviation (2), of Oconto River.

5-4 The Output

To show the application of the procedure for determining the reservoir capacity for any deterministic output, the periodic output is used in such a way that its first harmonic lags behind the first harmonic of the periodic input mean by $\pi/2$, so as to obtain substantial value of the difference storage capacity. The output is assumed to have two harmonics. Its plot is given in Fig. 5-2. However, output with any number of harmonics and with any phase difference can also be dealt with in a similar fashion.

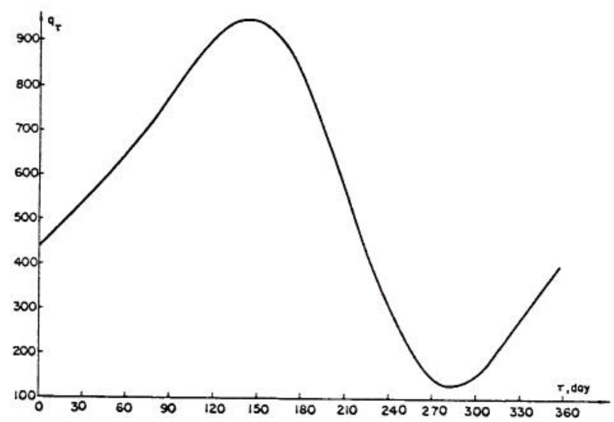

Fig. 5-2. Variation of Output Within the Year with Mean = 540, Fourier Coefficients A_1 = -187.21, B_1 = 324.24, A_2 = 79.00, and B_2 = -46.10.

5-5 Case Study

The simulated problem is to determine the storage capacity of a reservoir with the economic life of 100 years, to be constructed at the gauging site of the Oconto River at station number 4.0710 with the mean annual daily flow of 543.498 cfs.

5-6 Computation of the Difference Storage Capacity

The first problem is to finally estimate the periodic means, $\hat{\mu}_\tau$, and the periodic standard deviation, $\hat{\sigma}_\tau$, from the actual river flow data. Either a non-parameteric or a parameteric approach can be used. In this study a parameteric approach is used for the following reasons:

(1) In case of daily series, the number of parameters to be estimated for the mean and standard deviation is 730. The total number of statistics in the non-parameteric approach is so large that it becomes unfeasible to estimate accurately so many parameters from a limited sample size, with the estimates subject to large sampling errors.

(2) The non-parameteric method also removes all sampling variations associated with the mean and standard deviation from the stochastic component, which may represent often a large portion of the variance of the stochastic component.

In the harmonic analysis a maximum of six harmonics were fitted. However, only five harmonics were considered as representing the periodic mean and standard deviation, because the sixth harmonic gave little contribution to the explained variance of fluctuation of either the mean or the standard deviation, with the results given in Tables 5-1 and 5-2.

Table 5-1. Results of the Harmonic Analysis of the Mean Daily Values of the Oconto River

Harmonic	A	B	Explained Variance
1	- 200.3	- 112.4	0.3784
2	145.4	185.0	0.3971
3	- 85.5	- 79.9	0.0982
4	58.0	65.6	0.0550
5	- 39.8	- 72.5	0.0491
6	7.4	27.8	0.0059
Variance Explained by Six Harmonics			0.9837

Table 5-2. Results of the Harmonic Analysis of Daily Standard Deviations of the Oconto River

Harmonic	A	B	Explained Variance
1	- 123.3	- 85.6	0.2706
2	141.6	105.7	0.3750
3	- 66.4	- 46.2	0.0786
4	75.7	31.7	0.0809
5	- 47.2	- 43.2	0.0492
6	8.6	4.3	0.0011
Variance Explained by Six Harmonics			0.8554

The periodicity of daily flows is removed to give the y_t-series, as $y_t = (x_{p,\tau} - \hat{\mu}_\tau)/\hat{\sigma}_\tau$. This series is then standardized by using its mean $\bar{y} = 0.034$ and its standard deviation $s_y = 1.174$.

The resulting $\varepsilon_{p,\tau}$-series is then fitted by an autoregressive scheme, with the estimated coefficients given in Table 5-3. To satisfy the Wold's (1943) criteria, the $x^2 + 0.5418\ x + 0.3193 = 0$, with $x = (0.2709 + 0.4959\ i)$. The roots of x lie in a unit circle of the complex plane, thereby satisfying the necessary condition for the fitting of an autoregressive model.

Table 5-3. Fitting the Second-Order Autoregressive Model to Standardized Stochastic Component of Daily Flow Data of Oconto River

a_1	a_2	Variance of Residual Series	
		Theoretical	Computed
0.5418	0.3193	0.32908	0.32908

The $\varepsilon_{p,\tau}$-series is whitened, with the correlogram of the resulting $\eta_{p,\tau}$ independent component plotted in Fig. 5-3. Though all the correlation coefficients are close to zero, its independence cannot be checked by using the test given by Anderson (1941), because of very narrow tolerance band due to large sample size, (14,600 values). Quimpo (1967) felt the same difficulty. He stated that in the application of Anderson's test for $\rho=0$, however, despite a value of $r_1 = 0.05$, the size of the statistical sample was such that even this was, according to the test, still significantly different from zero. The same difficulty was encountered in applying the test given by Quenouille (1949) to the first and second-order autoregressive schemes. Therefore, the variance of the independent component is computed and compared with the theoretical variance. When these two variances are the same, then the assumed autoregressive model is considered to fit well the $\varepsilon_{p,\tau}$, stochastic component of the daily flow data. The theoretical variance is

$$\text{Var}\ \eta = \text{Var}\ \varepsilon[1 - a_1^2 - a_2^2 - \frac{2\ a_1^2\ a_2}{(1-a_2)}] \quad , \tag{5-1}$$

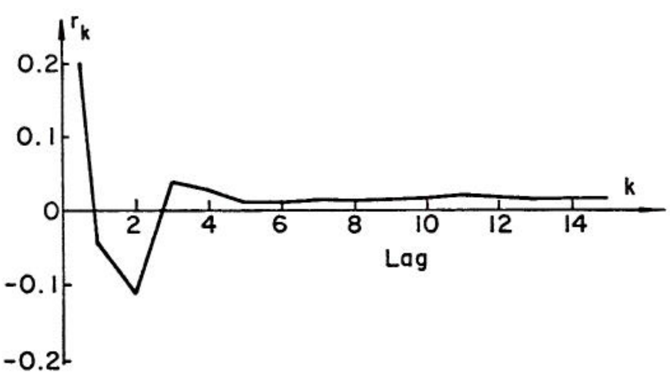

Fig. 5-3. Correlogram of Independent Component for Daily Flow Data of Oconto River.

where $\text{Var}\ \varepsilon = 1$ for the standardized $\varepsilon_{p,\tau}$-series. The daily flow data of this river is generated by Eq. (3-5) as

$$x_{p,\tau} = \hat{\mu}_\tau + \hat{\sigma}_\tau(0.034 + 1.174\ \varepsilon_{p,\tau}) \quad , \tag{5-2}$$

with

$$\varepsilon_{p,\tau} = 0.5418\ \varepsilon_{p,\tau-1} + 0.3193\ \varepsilon_{p,\tau-2} + 0.5736\ \eta_{p,\tau} \quad (5\text{-}3)$$

The mean maximum deficit for the output of Fig. 5-2 is found out from the generated net input series to give the value of estimated total storage. The stochastic storage is estimated by the mean maximum deficit of the generated stochastic net input series. The difference between the two storages gives the estimate of the difference storage tabulated in Table 5-4, and plotted in Fig. 5-4. It results from this figure that the difference storage oscillates around an approximate constant value of about 37,000. The difference storage capacity is then 37,000.

Table 5-4. Difference Storage for Periodic Output

n	Total Storage	Stochastic Storage	Difference Storage
2500	93624.672	56608.684	37015.988
3000	95006.483	56225.668	38780.815
3650	103444.409	65005.850	38438.559
4000	105752.440	68832.532	36919.908
4500	111133.409	73403.330	37730.079
5000	109970.113	73627.779	36342.334
5500	117266.204	80884.754	36381.450
6500	118464.232	82766.646	35697.586
7000	117821.429	79719.656	38101.773

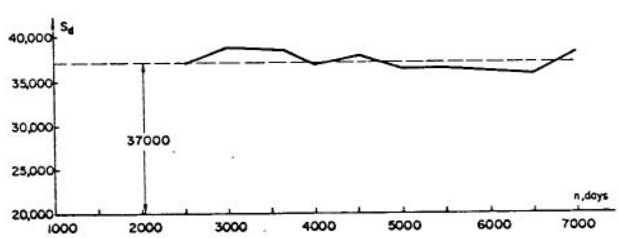

Figure 5-4. Difference Storage Capacity for Periodic Output of Fig. 5-2.

5-7 Determination of Stochastic Storage

The stochastic component of daily flow series has been fitted by second-order Markov model, namely

$$z_t = 0.5418\ z_{t-1} + 0.3193\ z_{t-2} + \eta_t \quad (5\text{-}4)$$

with $\sigma_{z,2} = 1.7432$. The model of Eq. (5-4) is equivalent to a first-order Markov model with $r_1 = 0.5418 + 0.3193 = 0.8611$. Therefore, $\sigma_{z,1} = 1.9668$. The stoastic storage was then worked out theoretically from Eq. (4-36), or from

$$f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \simeq C\sqrt{n}\ \left[\frac{\bar{\sigma}_\tau(1+\rho_1)^{1/2}\ \sigma_{z,1}}{(1-\rho_1)^{1/2}\ \sigma_{z,m}} - \bar{\sigma}_\tau + \sigma_n\right] \quad (5\text{-}5)$$

where $\bar{\sigma}_\tau = 288.37$ and $\hat{\sigma}_n = 344.56$. Substituting these values in Eq. (5-5), then

$$f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho] \simeq 1247.2158\ C\ \sqrt{n}\ . \quad (5\text{-}6)$$

The stochastic storage is obtained experimentally as the mean maximum deficit of the generated stochastic net input series. The stochastic net input series is $(x_{p,\tau} - \hat{\mu}_\tau)$, which is given by

$$(x_{p,\tau} - \hat{\mu}_\tau) = 0.034\ \hat{\sigma}_\tau + 1.174\ \hat{\sigma}_\tau\ \varepsilon_{p,\tau}\ . \quad (5\text{-}7)$$

The maximum deficit of $(0.034\ \hat{\sigma}_\tau)$ should be small as $\bar{y} = 0.034$ is small, while the expected maximum deficit of $1.174\ \hat{\sigma}_\tau\ \varepsilon_{p,\tau}$ is given by $s_y\ f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho]$ or $1.174\ f_4[\bar{\sigma}_\tau,\ s(\sigma_\tau),\ \rho]$. Thus the stochastic storage, S_s, is represented mathematically by

$$S_s \simeq 1247.2158\ C\ s_y\ \sqrt{n}\ ,$$

with

$$C \simeq \frac{S_s}{1464.23\sqrt{n}} \quad (5\text{-}8)$$

The C-values were calculated from the stochastic storage, obtained through generation of samples, needed in obtaining the difference storage capacity for different values of n by using Eq. (5-8). The values so obtained are plotted in Fig. 5-5.

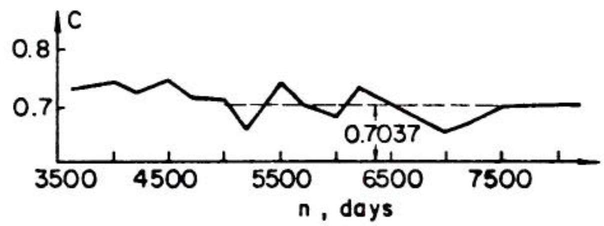

Fig. 5-5. Coefficient C Versus n.

It is found from Fig. 5-5 that initially and for smaller n, the value of C is high. The reason being that of a difference between the asymptotic and the exact value of the expected maximum deficit, because in Eq. (5-8) the asymptotic expression of expected maximum deficit was used. Hence, giving the more weight to values of C obtained for higher values of n, the average value of C is found to be 0.7037.

The value of the stochastic storage for a reservoir of 100 year economic life is then calculated for n = 36,500 by Eq. (5-8) as

$$S_s = 1464.\ 23\ C\ \sqrt{n} = 1464.\ 23 \times 0.7037 \times (36500)^{1/2}\ ,$$
$$= 196853\ .$$

5-8 Determination of the Total Storage

Since the total storage is the sum of difference storage capacity and the stochastic storage, its value is $S_t = S_d + S_s = 37000 + 196853.00 = 223853$. Thus the reservoir with 100 year economic life should be provided with a storage capacity of 223853 cfs-day. The reservoir capacity is thus 0.468 million acre ft.

5-9 Sensitivity Analysis for the Coefficient C

Since fixing the value of C by using Fig. 5-5 is subjective, it is considered worthwhile to study the effect of C on the error introduced in calculating the total storage. Let assume that the value of C for a particular n is 0.66 instead of 0.7037. This is almost the lowest value of C obtained in Fig. 5-5. Then $S_s = 1464.23 \times 0.66 \times (36500)^{1/2} = 184629.0580$, $S_t = 37000 + 184629.0580 = 221629$, or an error in design of 5.5%. This error is not too large, especially because it is the extreme case in Fig. 5-5. For all intermediate cases the errors would be smaller than 5.5%.

5-10 Sensitivity Analysis for the Difference Storage

Since fixing the value of difference storage capacity in Fig. 5-4 is also a subjective decision, it was considered worthwhile to study the effect of the error in the difference storage on the error introduced in calculating the total storage. Assume that for a particular n, the difference storage capacity is 35,000 instead of 37,000 as determined. Then S_s = 195853, S_d = 35,000, and S_t = 231853. The error is only 0.87%. A small error in difference storage capacity is not significant in the overall design of storage capacity of a reservoir.

Chapter 6
CONCLUSIONS

A methodology has been developed for designing the capacity of a *large* reservoir with a *high* level of development by using the concept of maximum deficit, and short-interval streamflow records, particularly daily flow data.

For this study, the input or the inflows were assumed to simulate daily river flows, which are composed of periodic and stochastic components. The stochastic components approximate Markov linear models, mainly second-order for daily flows. Output or the outflows were considered deterministic either as constants or as periodic functions.

From the analyses made for this study, the following conclusions were reached:

(1) The total storage capacity or the expected storage needed for the given output, when within-the-year fluctuation of the input is taken into consideration, can be divided into two parts: (a) a difference storage which is a function of periodic mean μ_τ and the periodic standard deviation σ_τ of the input, besides other parameters like amplitude, angular frequencies of various harmonics of input and output, and difference in phase between their harmonics; and (b) a stochastic storage, which is a function of the mean and standard deviation of σ_τ, of the autocorrelation coefficients of the Markov model considered, and of n;

(2) The difference storage stabilizes to a practically constant value beyond the stabilization region, which is inversely proportional to the level of development;

(3) The contribution of difference storage to the total storage or the needed capacity is significantly small in comparison to the stochastic storage, and hence a small error in its computation results hardly in a 1% error in the computed total storage for large reservoirs. Therefore, to save on computer time, it is felt that the difference storage capacity need not be computed very accurately;

(4) The asymptotic distribution of different statistics of partial sums for the dependent Markov models are functions of the asymptotic distribution of the respective statistics of the independent case; and

(5) The computation of the required storage capacity does not need the generation of flow sequences equal to the economic life of reservoir.

6-1 Recommendations for Further Research

There is some theoretical support for the stabilization of the difference storage for the case of full regulation. But, for partial regulation, Troutman (1976) showed mathematically that the total storage grows as $\ell n\ n$ for i.i.d. net inputs. The computer results of this study show that the difference storage becomes approximately constant for 90% development in case of a nonstationary independent input process, which could be because of $\sqrt{n}$ and $\ell n\ n$ curves being almost the same in the region of study. Hence, it is recommended that a further research should be undertaken for the case of partial regulation to ascertain mathematically how the total storage would grow with n for the case of a nonstationary net input process.

REFERENCES

Anderson, R.L., "Distribution of the Serial Correlation Coefficient," *Ann. Math. Statist.*, Vol. 13, pp. 1-13, 1941.

Anis, A.A., "The Variance of the Maximum Partial Sum of a Finite Number of Independent Normal Variates," *Biometrika*, Vol. 42, pp. 96-101,1955.

Anis, A.A., "On the Moments of the Maximum Partial Sums of a Finite Number of Independent Normal Variates," *Biometrika*, Vol. 43, pp. 79-84, 1956.

Anis, A.A., and Lloyd, E.H., "On the Range of Partial Sums of a Finite Number of Independent Normal Variates," *Biometrika*, Vol. 40, pp. 35-42, 1953.

Fathy, A., and Shukry, A.S., "The Problem of Reservoir Capacity for Long Term Storage," *J. Hydrol., Proc. Am. Soc. Civ. Engrs.*, Vol. 82, No. HY5, Paper 1082, pp. 1-27, 1956.

Feller, W., "The Asymptotic Distribution of the Range of Sums of Independent Random Variables," *Ann. Math. Statist.*, Vol. 22, pp. 427-432, 1951.

Fiering, M.B., "The Nature of the Storage Yield Relationship," *Proc. Symposium on Streamflow Regulation for Quality Control*, Public Health Service Publication No. 999-WP-30, pp. 243-253, U.S. Department of Health, Education and Welfare, Washington, D.C., 1965.

Fiering, M.B., "Stream Flow Synthesis," *Harvard University Press*, Cambridge, Massachusetts, 1967.

Gomide, F.L.S., "Range and Deficit Analysis Using Markov Chains," *Hydrology Paper No. 79*, Colorado State University, Fort Collins, Colorado, 1975.

Hardison, C.H., "Storage to Augment Low Flows," *Proc. Reservoir Yield Symposium*, Sept. 21-23, 1965, Part I, Paper 8, Discussion Part II, Water Research Association.

Hazen, A., "Storage to be Provided in Impounding Reservoirs for Municipal Water Supply," *Trans. Am. Soc. Civ. Engrs.*, Vol. 77, pp. 1539-1640, 1914.

Hurst, H.E., "Long-Term Storage Capacity of Reservoirs," *Trans. Am. Soc. Civ. Engrs.*, Vol. 116, pp. 770-779, 1951.

Lloyd, E.H., "Stochastic Reservoir Theory," *Adv. Hydrosci.*, Vol. 4, pp. 281-339, 1967.

Lloyd, E.H., "Wet and Dry Water," *Bulletin Institute of Mathematics and Its Applications*, Vol. 10, No. 9/10, pp. 348-353, 1974.

Lloyd, E.H., and Odoom, S., "Probability Theory of Reservoirs with Seasonal Input," *J. Hydrol.*, Vol. 2, pp. 1-10, 1964.

Melentijevich, M.J., "The Analysis of Range with Output Linearly Dependent Upon Storage," *Hydrology Papers*, Vol. 1, No. 11, Colorado State University, Fort Collins, Colorado, 1965.

Moran, P.A.P., "A Probability Theory of Dams and Storage Systems," *Aust. J. Appl, Sci.*, Vol. 5, pp. 116-124, 1954.

Moran, P.A.P., "A Probability Theory of Dams and Storage Systems: Modification of Release Rules," *Aust. J. Appl. Sci.*, Vol. 6, pp. 117-130, 1955.

Quenouille, M.H., "A Large Sample Test for the Goodness of Fit in Autoregressive Schemes," *J. Royal Statistical Soc.*, Vol. 110, pp. 123-129, 1949.

Quimpo, R.G., "Stochastic Model of Daily River Flow Sequences," *Hydrology Papers*, Vol. 1, No. 18, Colorado State University, Fort Collins, Colorado, 1967.

Rippl, W., "The Capacity of Storage Reservoirs for Water Supply," *Proc. Instn. Civ. Engrs.*, Vol. 71, pp. 270-278, 1883.

Roesner, L.A., and Yevjevich, V.M., "Mathematical Models for Time Series of Monthly Precipitation and Monthly Runoff," *Hydrology Papers*, Vol. 1, No. 15, Colorado State University, Fort Collins, Colorado, 1966.

Salas-La Cruz, J.D., "Range Analysis for Storage Problems of Periodic-Stochastic Processes," *Hydrology Papers*, Vol. 3, No. 57, Colorado State University, Fort Collins, Colorado, 1972.

Sudler, C.E., "Storage Required for the Regulation of Streamflow," *Trans. Am. Soc. Civ. Engrs.*, Vol. 91, pp. 622-660, 1927.

Sutabutra, P., "Reservoir Storage Capacity Required When Water Inflow Has a Periodic and a Stochastic Component," *Ph.D. Thesis*, Colorado State University, Fort Collins, Colorado, 1967.

Thomas, H.A., and Fiering, M.B., "Mathematical Synthesis of Stream Flow Sequences for the Analysis of River Basins by Simulation," *Design of Water Resources Systems*, Harvard University Press, Cambridge, Massachusetts, 1962.

Thomas, H.A., and Fiering, M.B., "The Nature of the Storage Yield Function," *Operations Research in Water Quality Management*, Harvard University Water Program, Chapter II, 1963.

Troutman, B.M., "Limiting Distributions in Storage Theory," *Ph.D. Thesis*, Colorado State University, Fort Collins, Colorado 1976.

Wold, H., "A Study in the Analysis of Stationary Time Series," *Uppsala Almquist and Wiksell*, 1954.

Yevjevich, V.M., "Fluctuations of Wet and Dry Years, Part II: Analysis by Serial Correlation," *Hydrology Papers*, Vol. 1, No. 4, Colorado State University, Fort Collins, Colorado, 1964.

Yevjevich, V.M., "The Application of Surplus, Deficit and Range in Hydrology," *Hydrology Papers*, Vol. 1, No. 10, Colorado State University, Fort Collins, Colorado, 1965.

Yevjevich, V.M., "Stochastic Problems in Design of Reservoir," *Water Resources Conference 7th*, Colorado State University, 1965, Published in *Water Research*, pp. 375-411, 1966.

Yevjevich, V.M., "Mean Range of Linearly Dependent Normal Variables with Applications to Storage Problems," *Water Resour. Res.*, Vol. 3, No. 3, pp. 663-671, 1967.

Yevjevich, V.M., "The Structure of Inputs and Outputs of Hydrologic Systems": *United States-Japan Bilateral Seminar in Hydrology*, Honolulu.

Yevjevich, V.M., "Stochastic Processes in Hydrology," *Water Resources Publications*, Fort Collins, Colorado, 1972.

Yevjevich, V.M., "Structural Analysis of Hydrologic Time Series," *Hydrology Papers*, Vol. 3, No. 56, Colorado State University, Fort Collins, Colorado, pp. 1-59, 1972a.

KEY WORDS: Water Storage, Flow Regulation, Deficit of Reservoir Capacity:

ABSTRACT: A methodology is presented for designing capacities of *large* reservoirs with sufficiently *high* levels of development by using the concept of maximum deficit rather than the range, and short-interval flow records of daily flows, composed of periodic and stochastic components. The output is assumed to be a deterministic process, either constant or periodic.

Analytical and data generation methods are used in computing the required storage capacities for regulating the periodic-stochastic inputs to deterministic outputs. The total storage as the mean required storage capacity is divided into the difference storage and the stochastic storage. Difference storage is due to periodic means of input and output. It is approximately constant for large sample sizes. It can be estimated for different sample sizes by generating a relatively short series of daily flows.

KEY WORDS: Water Storage, Flow Regulation, Deficit of Reservoir Capacity:

ABSTRACT: A methodology is presented for designing capacities of *large* reservoirs with sufficiently *high* levels of development by using the concept of maximum deficit rather than the range, and short-interval flow records of daily flows, composed of periodic and stochastic components. The output is assumed to be a deterministic process, either constant or periodic.

Analytical and data generation methods are used in computing the required storage capacities for regulating the periodic-stochastic inputs to deterministic outputs. The total storage as the mean required storage capacity is divided into the difference storage and the stochastic storage. Difference storage is due to periodic means of input and output. It is approximately constant for large sample sizes. It can be estimated for different sample sizes by generating a relatively short series of daily flows.

KEY WORDS: Water Storage, Flow Regulation, Deficit of Reservoir Capacity:

ABSTRACT: A methodology is presented for designing capacities of *large* reservoirs with sufficiently *high* levels of development by using the concept of maximum deficit rather than the range, and short-interval flow records of daily flows, composed of periodic and stochastic components. The output is assumed to be a deterministic process, either constant or periodic.

Analytical and data generation methods are used in computing the required storage capacities for regulating the periodic-stochastic inputs to deterministic outputs. The total storage as the mean required storage capacity is divided into the difference storage and the stochastic storage. Difference storage is due to periodic means of input and output. It is approximately constant for large sample sizes. It can be estimated for different sample sizes by generating a relatively short series of daily flows.

KEY WORDS: Water Storage, Flow Regulation, Deficit of Reservoir Capacity:

ABSTRACT: A methodology is presented for designing capacities of *large* reservoirs with sufficiently *high* levels of development by using the concept of maximum deficit rather than the range, and short-interval flow records of daily flows, composed of periodic and stochastic components. The output is assumed to be a deterministic process, either constant or periodic.

Analytical and data generation methods are used in computing the required storage capacities for regulating the periodic-stochastic inputs to deterministic outputs. The total storage as the mean required storage capacity is divided into the difference storage and the stochastic storage. Difference storage is due to periodic means of input and output. It is approximately constant for large sample sizes. It can be estimated for different sample sizes by generating a relatively short series of daily flows.

Stochastic storage, as the expected maximum deficit of stochastic net input, is determined analytically by using a coefficient, estimated by generating short series of daily flows. Parameters of asymptotic distributions of statistics of partial sums can be obtained by equating the higher-order autoregressive models to an equivalent first-order autoregressive model with ρ_1 equal to the sum of all their autoregressive coefficients.

The methodology is applied to a simulated problem of determining the storage capacity of a reservoir with the economic life of 100 years to be constructed at a site for which there are 40 years of daily flow data (the Oconto River near Gillett, Wisconsin).

REFERENCE: Mutreja, Kedar Nath, Colorado State University, Hydrology Paper No. 86 (September 1976) "Reservoir Capacity for Periodic Stochastic Input and Periodic Output."

Stochastic storage, as the expected maximum deficit of stochastic net input, is determined analytically by using a coefficient, estimated by generating short series of daily flows. Parameters of asymptotic distributions of statistics of partial sums can be obtained by equating the higher-order autoregressive models to an equivalent first-order autoregressive model with ρ_1 equal to the sum of all their autoregressive coefficients.

The methodology is applied to a simulated problem of determining the storage capacity of a reservoir with the economic life of 100 years to be constructed at a site for which there are 40 years of daily flow data (the Oconto River near Gillett, Wisconsin).

REFERENCE: Mutreja, Kedar Nath, Colorado State University, Hydrology Paper No. 86 (September 1976) "Reservoir Capacity for Periodic Stochastic Input and Periodic Output."

Stochastic storage, as the expected maximum deficit of stochastic net input, is determined analytically by using a coefficient, estimated by generating short series of daily flows. Parameters of asymptotic distributions of statistics of partial sums can be obtained by equating the higher-order autoregressive models to an equivalent first-order autoregressive model with ρ_1 equal to the sum of all their autoregressive coefficients.

The methodology is applied to a simulated problem of determining the storage capacity of a reservoir with the economic life of 100 years to be constructed at a site for which there are 40 years of daily flow data (the Oconto River near Gillett, Wisconsin).

REFERENCE: Mutreja, Kedar Nath, Colorado State University, Hydrology Paper No. 86 (September 1976) "Reservoir Capacity for Periodic Stochastic Input and Periodic Output."

Stochastic storage, as the expected maximum deficit of stochastic net input, is determined analytically by using a coefficient, estimated by generating short series of daily flows. Parameters of asymptotic distributions of statistics of partial sums can be obtained by equating the higher-order autoregressive models to an equivalent first-order autoregressive model with ρ_1 equal to the sum of all their autoregressive coefficients.

The methodology is applied to a simulated problem of determining the storage capacity of a reservoir with the economic life of 100 years to be constructed at a site for which there are 40 years of daily flow data (the Oconto River near Gillett, Wisconsin).

REFERENCE: Mutreja, Kedar Nath, Colorado State University, Hydrology Paper No. 86 (September 1976) "Reservoir Capacity for Periodic Stochastic Input and Periodic Output."

LIST OF PREVIOUS 25 PAPERS

No. 61 Drainage Design Based Upon Aeration by Harold R. Duke, June 1973.

No. 62 Techniques for Modeling Reservoir Salinity by John Hendrick, August 1973.

No. 63 Mechanics of Soil Erosion From Overland Flow Generated by Simulated Rainfall by Mustafa Kilinc and Everett V. Richardson, September 1973.

No. 64 Area-Time Structure of the Monthly Precipitation Process by V. Yevjevich and Alan K. Karplus, August 1973.

No. 65 Almost-Periodic, Stochastic Process of Long-Term Climatic Changes, by William Q. Chin and Vujica Yevjevich, March 1974.

No. 66 Hydrologic Effects of Patch Cutting of Lodgepole Pine, by Thomas L. Dietrich and James R. Meiman, April 1974.

No. 67 Economic Value of Sediment Discharge Data, by Sven Jacobi and Everett V. Richardson, April 1974.

No. 68 Stochastic Analysis of Groundwater Level Time Series in the Western United States, by Albert G. Law, May 1974.

No. 69 Efficient Sequential Optimization in Water Resources, by Thomas E. Croley II, September 1974.

No. 70 Regional Water Exchange for Drought Alleviation, by Kuniyoshi Takeuchi, November 1974.

No. 71 Determination of Urban Watershed Response Time, by E. F. Schulz, December, 1974.

No. 72 Generation of Hydrologic Samples, Case Study of the Great Lakes, by V. Yevjevich May, 1975.

No. 73 Extraction of Information on Inorganic Water Quality, by William L. Lane, August, 1975.

No. 74. Numerical Model of Flow in Stream-Aquifer System, by Catherine E. Kraeger Rovey, August, 1975.

No. 75. Dispersion of Mass in Open-Channel Flow, by William W. Sayre, August, 1975.

No. 76 Analysis and Synthesis of Flood Control Measures, by Kon Chin Tai, September, 1975.

No. 77 Methodology for the Selection and Timing of Water Resources Projects to Promote National Economic Development, by Wendim-Agegnehu Lemma, August 1975.

No. 78 Two-Dimensional Mass Dispersion in Rivers, by Forrest M. Holly, Jr., September 1975.

No 79. Range and Deficit Analysis Using Markov Chains, by Francisco Gomide, October, 1975.

No. 80 Analysis of Drought Characteristics by the Theory of Run, by Pedro Guerrero-Salazar and Vujica Yevjevich, October 1975.

No. 81 Influence of Simplifications in Watershed Geometry in Simulation of Surface Runoff, by L. J. Lane, D. A. Woolhiser and V. Yevjevich, January 1976.

No. 82 Distributions of Hydrologic Independent Stochastic Components, by Pen-chih Tao, V. Yevjevich and N. Kottegoda, January 1976.

No. 83 Optimal Operation of Physically Coupled Surface and Underground Storage Capacities, by Dragoslav Isailović, January 1976.

No. 84 A Salinity Management Strategy for Stream-Aquifer Systems, by Otto J. Helweg, and J. Labadie, February 1976.

No. 85 Urban Drainage and Flood Control Projects Economic, Legal and Financial Aspects, by Neil S. Grigg, Leslie H. Botham, Leonard Rice, W. J. Shoemaker, and L. Scott Tucker, February 1976.