

Error Analysis of Conventional Discrete and Gradient Dynamic Programming

PETER K. KITANIDIS AND EFI FOUFOULA-GEORGIU¹

St. Anthony Falls Hydraulic Laboratory, Department of Civil and Mineral Engineering, University of Minnesota, Minneapolis

An asymptotic error analysis of the conventional discrete dynamic programming (DDP) method is presented, and upper bounds of the error in the control policy (i.e., the difference of the estimated and true optimal control) at each operation period are computed. This error is shown to be of the order of the state discretization interval (ΔS), a result with significant implications in the optimization of multi-state systems where the "curse of dimensionality" restricts the number of states to a relatively small number. The error in the optimal cost varies with ΔS^2 . The analysis provides useful insights into the effects of state discretization on calculated control and cost functions, the comparability of results from different discretizations, and criteria about the required number of nodes. In an effort to reduce the discretization error in the case of smooth cost functions, a new discrete dynamic programming method, termed gradient dynamic programming (GDP), is proposed. GDP uses a piecewise Hermite interpolation of the cost-to-go function, at each stage, which preserves the values of the cost-to-go function and of its first derivatives at the discretization nodes. The error in the control policy is shown to be of the order of $(\Delta S)^3$ and the error in the cost to vary with ΔS^4 . Thus as ΔS decreases, GDP converges to the true optimum much more rapidly than DDP. Another major advantage of the new methodology is that it facilitates the use of Newton-type iterative methods in the solution of the nonlinear optimization problems at each stage. The linear convergence of DDP and the superlinear convergence of GDP are illustrated in an example.

1. INTRODUCTION

Dynamic programming has found many applications in water resources planning, particularly in the optimization of reservoir operations. In fact, water resource problems have served as a stimulus to the development of dynamic programming and a water resources study by Massé [1946] preceded in publication the foundational work of Bellman [1952] in advancing the functional equations of dynamic programming.

The literature on dynamic programming (or "multistage optimization") theory and applications is very extensive, the books of Bellman [1957] and Bellman and Dreyfus [1962] being the classic references. Dynamic programming solves the overall optimization problem in stages and, whenever applicable, reduces the cost of computations. In general, if the problem can be decomposed into N stages with m decision variables at each stage, the cost increases approximately with $N m^2$, as compared with $N^2 m^2$ for solving the problem in a single stage. State of the art reviews with extensive lists of references on dynamic programming and its application are to be found in the work by Yakowitz [1982] for several water resource problems, Yeh [1985] for optimal reservoir operation, and Kitanidis [1983] for real-time optimal reservoir operation. In this work we will assume that the reader is familiar with the basic concepts of dynamic programming, such as decomposition into stages, state variables, and the cost-to-go function, so that we may concentrate on computational aspects of discrete dynamic programming.

Discrete dynamic programming (DDP) has been the most common numerical method used for the application of dynamic programming. First proposed by Bellman [1957], it has found numerous applications in water resources planning: Buras [1963], Gablinger and Loucks [1970], Loucks and Falkson [1970], Butcher [1971], Su and Deininger [1974], and

Askew [1974], to mention a few of the earliest works. However, in the past 10 years considerable effort has been directed toward developing new computational techniques for the solution of the functional equations of dynamic programming. The objective of these methods is to avoid the main computational disadvantage of DDP, known as the "curse of dimensionality" (see, for example, Yakowitz [1982] for a discussion of these methods and appropriate references). They are iterative schemes that concentrate on obtaining an optimal state-control trajectory, starting from a given initial state and subject to deterministic inputs. In this case one can do away with the discretization of state variables (under lenient conditions of differentiability) and the optimization is with respect to a finite number of decision variables. Such methods, which furthermore make use of efficient nonlinear programming techniques to optimize at each stage, can be much faster than DDP.

It must be realized, however, that DDP was designed to solve a more difficult problem than obtaining a single optimal trajectory: to find at each stage the optimal control policy, i.e., the control variables as functions of the continuous state variables, so that a given criterion is optimized. The calculation of the optimal policy at all stages is desirable in many cases and required in stochastic systems where no single trajectory can be projected with certainty. The difficulty is that now, optimization is with respect to a finite number of decision functions (minimization of a cost functional). Unless analytical solutions are available, as in the linear quadratic Gaussian case [Wasimi and Kitanidis, 1983; Laoiciga and Mariño, 1985] or in small-perturbation approximations [Kitanidis, 1985], one must resort to computer-based numerical methods which approximate optimization of a functional (an "infinite-dimensional" problem) by optimization with respect to a finite number of decision variables, a problem amenable to computer-based methods of solution. This approximation introduces a discretization error, such as when a function is substituted by its values at some points, and is also the root of the dimensionality curse. Let it be emphasized that neither the discretization error nor the dimensionality curse are innate to dynamic programming but plague whatever method of solu-

¹Now at Department of Civil Engineering, University of Iowa, Ames, Iowa.

Copyright 1987 by the American Geophysical Union.

tion substitutes a decision function with a finite number of decision variables.

In computing the optimal control policy of a continuous state-space system via DDP, cost-to-go functions of continuous state variables are replaced by sets of values at discrete grid points. Values at points between nodes are obtained through interpolation. It is understood that the solution of the resulting approximate problem (estimated control policy) will be different from the solution of the original problem (true control policy). Solution accuracy is an important but seldom considered aspect of discrete dynamic programming.

Yakowitz [1982] cautions that "one should be wary of proposed solutions until the deleterious effects of discretization have been somehow bounded and found acceptable." From a practical viewpoint, the question is, How fine a discretization is needed? According to Klemeš [1977a], Savarenskiy [1940] and Doran [1975] recommended discretization of each state variable into 5 to 10 nodes and Moran [1959] into 15 to 20. These authors were apparently concerned with the accuracy in calculating probabilities (e.g., calculation of probability of full reservoir). Klemeš [1977a, p. 149] performed theoretical and numerical studies which showed that too coarse a discrete storage representation "may completely distort reality in most unexpected ways.... An inadequate number of storage states, besides causing a decrease in accuracy, may result in a gradual collapse of the optimization scheme. This collapse may escape attention, since the computer algorithm may keep working and producing results that may even seem reasonable."

In this paper, an asymptotic error analysis of the conventional discrete dynamic programming method is presented, and upper bounds on the error in the control policy (the difference of the estimated and true control policies) at each operation period are computed. This error is shown to be of the order of the state discretization interval ΔS , a result with significant practical implications, especially in multistate systems where the curse of dimensionality restricts the number of discretization nodes for each state variable to a relatively small number. The error in the estimation of the minimum cost is shown to vary with ΔS^2 .

Another contribution of this paper is a new discrete dynamic programming method, which is termed gradient dynamic programming (GDP). This method is based on a piecewise Hermite interpolation of the cost-to-go function at each stage. In this method the values of the cost to go and the values of its first derivatives (gradient) are calculated on the nodes of the grid and preserved. The GDP method is shown to yield an estimated control policy whose asymptotic proximity to the true control policy varies with $(\Delta S)^3$, while the estimate of the optimum cost varies with ΔS^4 . This implies that halving the state discretization interval improves the solution of DDP by a factor of 2, whereas the solution of GDP is improved by a factor of 8. That the above asymptotic results hold well even for finite ΔS is demonstrated in an example involving the optimal control of a single-reservoir system.

2. ASYMPTOTIC ERROR ANALYSIS OF CONVENTIONAL DISCRETE DYNAMIC PROGRAMMING

For the sake of simplicity our analysis will be limited to the scalar deterministic case. Analysis for the multivariate case would follow the same lines and the conclusions would not be affected. We will discuss this issue as well as the effect of stochastic inputs at the end of the section.

Consider a single-reservoir system to be operated over a time horizon of N periods ("stages"). Let $S(k)$ denote the reservoir storage (state variable) at the end of period k , and $u(k + 1)$ and $q(k + 1)$ denote the regulated release (control variable) and inflow, respectively, during period $k + 1$. The constraints on the system consist of the continuity equation, the nonnegativity and capacity constraints on the storage, and the nonnegativity constraint on the control:

$$S(k + 1) = S(k) + q(k + 1) - u(k + 1) \tag{1a}$$

$$0 \leq S(k + 1) \leq K \tag{1b}$$

$$0 \leq u(k + 1) \tag{1c}$$

Let $F_k[S(k)]$ denote the cost-to-go function when there are $N - k$ periods to go. Under the usual assumptions of multistage optimization, the functional equation of the system can be written as

$$F_k[S(k)] = \min_{u(k+1)} \{c_k[S(k), u(k + 1)] + F_{k+1}[S(k + 1)]\} \tag{2}$$

$k = N - 1, \dots, 0$

where $F_N[S(N)]$ is a given function of the final storage, and $c_k(\)$ is the loss function at stage k , a function of the storage at the beginning of period $k + 1$, and the release during period $k + 1$. For the error analysis which is performed in this section, it will be required to assume that these functions are three times differentiable, although the third derivative does not need to be continuous. If $u^*(k + 1)$ denotes the optimal release corresponding to $S(k)$, (2) yields a recursive equation for the computation of the cost-to-go function:

$$F_k[S(k)] = c_k[S(k), u^*(k + 1)] + F_{k+1}[S(k) + q(k + 1) - u^*(k + 1)] \tag{3}$$

If one of the constraints (1b) or (1c) is binding, it uniquely determines the solution. In our case the discretization error affects the control when no constraint is binding. Omitting the arguments of $c_k(\)$ and $F^{k+1}(\)$, the true optimal release $U^*(k + 1)$ is the solution to equation

$$\partial c_k / \partial u + (dF_{k+1}) / du = 0 \tag{4a}$$

or

$$\partial c_k / \partial u - (dF_{k+1}) / dS = 0 \tag{4b}$$

Note that the minus sign appeared from the application of the chain rule of differentiation, i.e.,

$$\frac{dF_{k+1}}{du(k + 1)} = \frac{dF_{k+1}}{dS(k + 1)} \frac{dS_{(k+1)}}{du(k + 1)} = - \frac{dF_{k+1}}{dS(k + 1)}$$

In many applications of DDP the search for optimum is restricted to the finite set of $u(k + 1)$ for which $S(k) + q(k + 1) - u(k + 1)$ is a node. If this is the case, it is obvious that the error in the control is of order ΔS , the state discretization interval. For differentiable c_k and F_{k+1} , a potentially more accurate procedure is to include in the search the solution to an approximation of (4b) in which the cost-to-go function is replaced by an interpolation scheme which uses the nodal values. Since such issues are seldom discussed in the DDP literature, our interpretation is that the cost-to-go function is approximated by the simplest possible scheme, i.e., a piecewise linear function which reproduces the values calculated at the nodes. Let S_{l-1} and S_l denote the lower and upper

grid points (storage states) between which the storage $S(k) + q(k + 1) - u^*(k + 1)$ falls, where u^* is the calculated control. Then (4b) is approximated in conventional discrete dynamic programming through

$$\frac{\partial c_k}{\partial u} - \frac{F_{k+1}'(S_i) - F_{k+1}'(S_{i-1})}{\Delta S} = 0 \tag{5}$$

where $F_{k+1}'(S)$ is the piecewise linear function which reproduces the cost to go computed at the nodes, and ΔS is the storage discretization interval; $u^*(k + 1)$ is determined from the solution of (5), which solution is assumed to exist. (That is, we will now deal with the case that $u^*(k + 1)$ does not correspond to a node.) Let $F_{k+1}'(S) = F_{k+1}(S) + P(S)$. The estimated optimal release $u^*(k + 1)$ differs from the true optimal release $U^*(k + 1)$ by Δu_{k+1} ,

$$u^*(k + 1) = U^*(k + 1) + \Delta u_{k+1} \tag{6}$$

Our intention is to relate Δu_{k+1} to ΔS and, in particular, to determine the order of convergence of $u^*(k + 1)$ to $U^*(k + 1)$ through asymptotic analysis [see Luenberger, 1973, p. 127].

A Taylor series expansion of the terms of (5) around $U^*(k + 1)$ gives

$$\begin{aligned} &\frac{\partial c_k}{\partial u} \Big|_{U^*(k+1)} + \frac{\partial^2 c_k}{\partial u^2} \Big|_{U^*(k+1)} \Delta u_{k+1} - \frac{1}{\Delta S} \left\{ F_{k+1} \Big|_{A^*} \right. \\ &+ \frac{dF_{k+1}}{dS} \Big|_{A^*} (S_i - A^*) + \frac{1}{2} \frac{d^2 F_{k+1}}{dS^2} \Big|_{A^*} (S_i - A^*)^2 - F_{k+1} \Big|_{A^*} \\ &\left. - \frac{dF_{k+1}}{dS} \Big|_{A^*} (S_{i-1} - A^*) - \frac{1}{2} \frac{d^2 F_{k+1}}{dS^2} \Big|_{A^*} (S_{i-1} - A^*)^2 \right\} \\ &- \frac{1}{\Delta S} [P_{k+1}(S_i) - P_{k+1}(S_{i-1})] \\ &+ O(\Delta u_{k+1}^2) + O(\Delta S^2) = 0 \end{aligned} \tag{7}$$

where the symbol $A^* \equiv S(k) + q(k + 1) - U^*(k + 1)$ has been introduced for convenience in notation, and O represents the order of terms which have been neglected. These orders were obtained by observing that the lowest-order omitted terms are

$$\begin{aligned} &\frac{1}{2} \frac{\partial^3 c_k}{\partial u^3} \Delta u_{k+1}^2 + \frac{1}{6} \frac{1}{\Delta S} \frac{d^3 F_{k+1}}{dS^3} [(S_i - A^*)^3 - (S_{i-1} - A^*)^3] \\ &= \frac{1}{2} \frac{\partial^3 c_{k+1}}{\partial u^3} \Delta u_{k+1}^2 + \frac{1}{6} \frac{d^3 F_{k+1}}{dS^3} [(S_i - A^*)^2 \\ &+ (S_i - A^*)(S_{i-1} - A^*) + (S_{i-1} - A^*)^2] \end{aligned}$$

Then setting $B_i = S_i - [S(k) + q(k + 1) - u^*(k + 1)]$, which is of order ΔS ,

$$\begin{aligned} (S_i - A^*)^2 &= (B_i - \Delta u_{k+1})^2 \\ &= B_i^2 + \Delta u_{k+1}^2 - 2B_i \Delta u_{k+1} = O(\Delta u_{k+1}^2) + O(\Delta S^2) \end{aligned}$$

and similarly for $(S_{i-1} - A)$.

Returning to (7), since $S_i - S_{i-1} = \Delta S$, the above equation

reduces to

$$\begin{aligned} &\frac{\partial c_k}{\partial u} \Big|_{U^*(k+1)} - \frac{dF_{k+1}}{dS} \Big|_{A^*} + \frac{\partial^2 c_k}{\partial u^2} \Big|_{U^*(k+1)} \Delta u_{k+1} \\ &- \frac{d^2 F_{k+1}}{dS^2} \Big|_{A^*} \left(\frac{S_i + S_{i-1}}{2} - A^* \right) - \frac{1}{\Delta S} [P(S_i) - P(S_{i-1})] \\ &+ O(\Delta u_{k+1}^2) + O(\Delta S^2) = 0 \end{aligned} \tag{8}$$

Since $U^*(k + 1)$ is the solution of (4b), the sum of the first two terms vanishes and, after simplifying the notation, (8) further reduces to

$$\begin{aligned} &\frac{\partial^2 c_k}{\partial u^2} \Delta u_{k+1} - \frac{d^2 F_{k+1}}{dS^2} \left\{ \frac{S_i + S_{i-1}}{2} - [S(k) + q(k + 1) - U^*(k + 1)] \right\} \\ &- \frac{1}{\Delta S} [P_{k+1}(S_i) - P_{k+1}(S_{i-1})] = 0 \end{aligned} \tag{9}$$

Denote by ε the distance of the midpoint of the interval (S_{i-1}, S_i) to the storage value $S(k) + q(k + 1) - u^*(k + 1)$; i.e., $\varepsilon = (S_i + S_{i-1})/2 - [S(k) + q(k + 1) - u^*(k + 1)]$ (see Figure 1). Then (9) can be simplified to

$$\begin{aligned} &\frac{\partial^2 c_k}{\partial u^2} \Delta u_{k+1} + \frac{d^2 F_{k+1}}{dS^2} (\Delta u_{k+1} - \varepsilon) \\ &- \frac{1}{\Delta S} [P_{k+1}(S_i) - P_{k+1}(S_{i-1})] = 0 \end{aligned} \tag{10}$$

Assuming that $U^*(k + 1)$ is a unique local minimum so that the second derivative of the objective function is positive, i.e.,

$$\partial^2 c_k / \partial u^2 + (d^2 F_{k+1}) / dS^2 > 0$$

(10) can be solved in terms of Δu_{k+1}

$$\begin{aligned} \Delta u_{k+1} &= \frac{(d^2 F_{k+1}) / dS^2}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \varepsilon + \frac{1}{\Delta S} \frac{1}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \\ &\cdot [P_{k+1}(S_i) - P_{k+1}(S_{i-1})] \end{aligned} \tag{11}$$

where all derivatives are calculated at values corresponding to U^* .

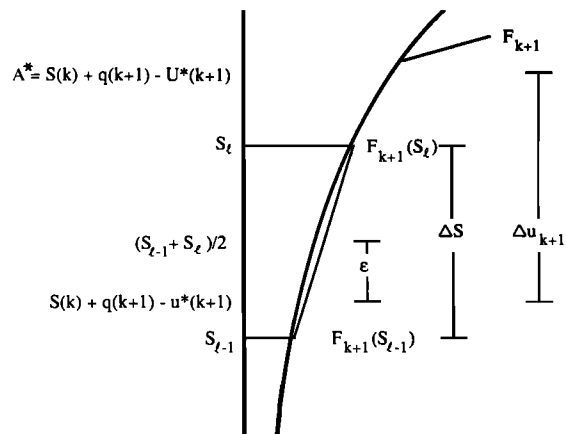


Fig. 1. Schematic representation of the cost to go function at stage $k + 1$ and definition of some terms used in the analysis.

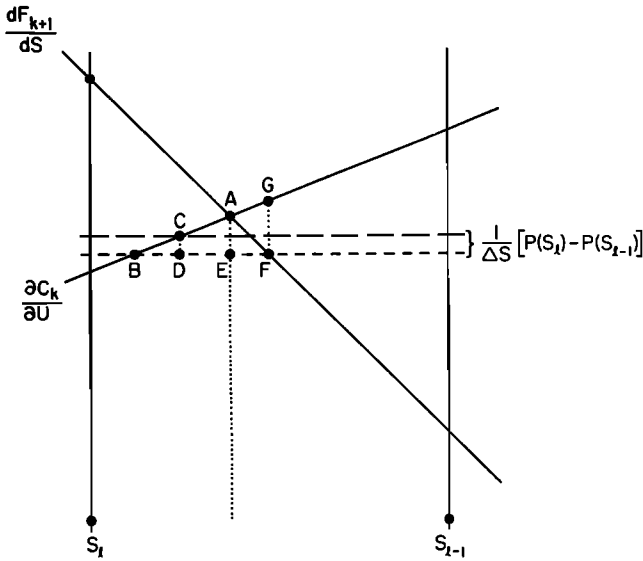


Fig. 2. Geometric representation of the error in the estimation of optimal control.

The solution is also given geometrically in Figure 2. The two solid straight lines represent $(dF_{k+1})/dS$ and $\partial c_k/\partial u$ plotted against $S(k+1)$ (for given initial value $S(k)$ and input $q(k+1)$). These lines are depicted as straight because neglecting higher-order terms is equivalent to assuming that F_{k+1} and c_k are quadratic at the scale of ΔS . Point A is the point which satisfies (4b) and consequently corresponds to the true optimum. Point B is the intersection of $\partial c_k/\partial u$ and the horizontal line at $[F_{k+1}(S_i) - F_{k+1}(S_{i-1})](\Delta S)^{-1}$. (Note that this horizontal line crosses dF_{k+1}/dS at F, the midpoint between S_i and S_{i-1} .) Consequently, B corresponds to the control calculated from (5) with the actual cost-to-go function F_{k+1} . Finally, C is the intersection of $\partial c_k/\partial u$ and the horizontal line at

$$[F_{k+1}(S_i) - F_{k+1}(S_{i-1})](\Delta S)^{-1} + [P_{k+1}(S_i) - P_{k+1}(S_{i-1})](\Delta S)^{-1}$$

Thus C is the point which corresponds to $u^*(k+1)$. The error Δu_{k+1} is depicted geometrically by segment DE and ϵ by segment FD. Then, $\Delta u_{k+1} = BE - BD$.

$$\frac{BE}{\epsilon + BD} = \frac{EA}{FG} = \frac{[(d^2 F_{k+1})/dS^2]}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}}$$

$$BD = \frac{\frac{1}{\Delta S} [P_{k+1}(S_i) - P_{k+1}(S_{i-1})]}{\partial^2 c_k/\partial u^2}$$

Combining, we obtain (11).

Now, we turn our attention to establishing a recursive equation for the calculation of $P_k(S)$. Given the notation established earlier, the cost to go may be computed in DDP using linear interpolation:

$$F_k[S(k)] = c_k[S(k), u^*(k+1)] + F_{k+1}(S_i)[S(k) + q(k+1) - u^*(k+1) - S_{i-1}](\Delta S)^{-1} + F_{k+1}(S_{i-1})[S_i - S(k) - q(k+1) + u^*(k+1)](\Delta S)^{-1} \quad (12)$$

while the truly minimum cost to go is given by

$$F_k[S(k)] = c_k[S(k), U^*(k+1)] + F_{k+1}[S(k+1) + q(k+1) - U^*(k+1)] \quad (13)$$

Expanding the terms at (12) around $U^*(k+1)$ and accounting for (4b) and the definition of $P(S)$ we obtain

$$P_k[S(k)] = P_{k+1}(S_i)[S(k) + q(k+1) - u^*(k+1) - S_{i-1}](\Delta S)^{-1} + P_{k+1}(S_{i-1})[S_i - S(k) - q(k+1) + u^*(k+1)](\Delta S)^{-1} + \frac{1}{2} \left\{ \frac{\partial^2 c_k}{\partial u^2} \Delta u_{k+1}^2 + \frac{d^2 F_{k+1}}{dS^2} \left[(S_i - A^*)^2 \frac{A^* - S_{i-1} - \Delta u_{k+1}}{\Delta S} + (S_{i-1} - A^*)^2 \frac{S - A^* + \Delta u_{k+1}}{\Delta S} \right] \right\} \quad (14)$$

where terms known to be of order ΔS^3 , Δu_{k+1}^3 are neglected. Using the definition of ϵ , (14) is now simplified to read

$$P_k[S(k)] = P_{k+1}(S_i)[S(k) + q(k+1) - u^*(k+1) - S_{i-1}](\Delta S)^{-1} + P_{k+1}(S_{i-1})[S_i - S(k) - q(k+1) + u^*(k+1)](\Delta S)^{-1} + \frac{1}{2} \left(\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2} \right) \Delta u_{k+1}^2 + \frac{1}{2} \frac{d^2 F_{k+1}}{dS^2} \left(\frac{\Delta S^2}{4} - \epsilon^2 \right) \quad (15)$$

where Δu_{k+1} is given by (11).

With (11) and (15) at our disposal we can now show that $\Delta u_{k+1} \sim O(\Delta S)$ and $P_k \sim O(\Delta S^2)$ at every $k = 1, \dots, N-1$. The proof will be given by mathematical induction. At the last stage, $k = N-1$, P_{k+1} vanishes, since the terminal cost function is given. Then taking absolute values of the terms of (11)

$$|\Delta u_N| = \frac{|d^2 F_N/dS^2|}{\frac{\partial^2 c_{N-1}}{\partial u^2} + \frac{d^2 F_N}{dS^2}} |\epsilon| \leq \frac{|d^2 F_N/dS^2|}{\frac{\partial^2 c_{N-1}}{\partial u^2} + \frac{d^2 F_N}{dS^2}} \frac{\Delta S}{2} \quad (16)$$

Thus Δu_N is of order ΔS . Taking absolute values of the terms of (15)

$$|P_{N-1}[S(N-1)]| \leq \frac{1}{8} \left[\frac{[(d^2 F_N)/dS^2][(\partial^2 c_{N-1})/\partial u^2]}{\frac{\partial^2 c_{N-1}}{\partial u^2} + \frac{d^2 F_N}{dS^2}} + \left| \frac{d^2 F_N}{dS^2} \right| \right] \Delta S^2 \quad (17)$$

Thus P_{N-1} is of order ΔS^2 . It remains to show that if $P_{k+1} \sim O(\Delta S^2)$, then $u_{k+1} \sim O(\Delta S)$ and $P_k \sim O(\Delta S^2)$. From (11),

$$|\Delta u_{k+1}| \leq \frac{|(d^2 F_{k+1})/dS^2|}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \frac{\Delta S}{2} + \frac{1}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \frac{|P_{k+1}(S_i) - P_{k+1}(S_{i-1})|}{\Delta S} \quad (18)$$

Both terms on the right-hand side are of order ΔS . In (15), all terms on the right-hand side are of order ΔS^2 . Taking absolute

values,

$$\begin{aligned}
 |P_k[S(k)]| \leq & \frac{1}{8} \left[\frac{|[(d^2 F_{k+1})/dS^2][(\partial^2 c_k/\partial u^2)]|}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} + \left| \frac{d^2 F_{k+1}}{dS^2} \right| \right] \Delta S^2 \\
 & + \frac{1}{2} \frac{1}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \{ [P_{k+1}(S_i) - P_{k+1}(S_{i-1})]^2 \} (\Delta S^2)^{-1} \\
 & \cdot \frac{1}{2} \frac{(d^2 F_{k+1})/dS^2}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} |P_{k+1}(S_i) - P_{k+1}(S_{i-1})| \\
 & + |P_{k+1}(S_i)| [S(k) + q(k+1) - u^*(k+1) - S_{i-1}] (\Delta S)^{-1} \\
 & + |P_{k+1}(S_{i-1})| [S_i - S(k) - q(k+1) + u^*(k+1)] (\Delta S)^{-1}
 \end{aligned} \tag{19}$$

Equations (18) and (19) are the sought after upper bounds on the error in the control and the cost-to-go functions, respectively.

Thus our analysis has shown that $\Delta u_k \sim O(\Delta S)$ and $|F_{k+1}' - F_{k+1}| \sim O(\Delta S^2)$. To avoid unnecessary complications, the proof was given for the scalar case. This is common practice wherever asymptotic analysis is used in optimization or numerical analysis. Exactly the same methodology can be used in the multivariate case using matrix-vector notation, the formulae for the expansion into Taylor series of functions of many variables, and retention of the lowest-order terms. As in the scalar case, it can be shown that as the discretization intervals decrease in size, DDP converge linearly in the control and quadratically in the cost.

In the stochastic case, the objective is to minimize the expected value of the expression of (2):

$$\begin{aligned}
 F_k[S(k)] = \min_{u(k+1)} \{ & c_k[S(k), u(k+1)] \\
 & + E_{q(k+1)} [F_{k+1}[S(k) + q - u(k+1)]] \}
 \end{aligned} \tag{20}$$

where

$$\begin{aligned}
 E_{q(k+1)} [F_{k+1}[S(k+1) + q - u(k+1)]] \\
 = \int_{\text{all } q} F_{k+1}[S(k+1) + q - u(k+1)] p(q) dq
 \end{aligned}$$

and $p(q)$ is the probability density function of inflow q . Then, one can repeat the same analysis except that the cost-to-go function is replaced by its expected value. The conclusion that $|\Delta u_{k+1}| \sim O(\Delta S)$ and $|F_{k+1}' - F_{k+1}| \sim O(\Delta S^2)$ would remain unaffected by stochasticity. However, the coefficients of proportionality would be affected since they would depend on the second derivative of $E[F_{k+1}]$ rather than of F_{k+1} . Expectation is weighted averaging which smooths out variability in F_{k+1} . For convex F_{k+1} , the maximum value of the second derivative of $E[F_{k+1}]$ is smaller than the maximum of the second derivative of F_{k+1} . Thus the maximum error for the stochastic case tends to be smaller than the maximum error for the deterministic case with the same cost functions, constraints, and the mean values of the inflows.

3. PRACTICAL SIGNIFICANCE OF THE RESULTS

The analysis was motivated from a study of the operation of the reservoirs of the Des Moines River, in Center, Iowa, using DDP [e.g., Zaphirakos, 1982; Wasimi and Kitanidis, 1983; Collado, 1984]. Like other practitioners before us, we were concerned about the effect of state (storage) discretization on the accuracy of the control (release) and cost of operation determined through DDP. A review of the literature revealed no lack of case studies which report on the number of nodes which was considered "adequate" in a particular problem. However, we could find no rules which indicate whether these same numbers would be applicable to another case study with different cost functions, constraints, and inputs.

Evaluating the effect of state discretization is particularly important in multistate problems where computational considerations often preclude the use of fine grids. It is also important because, as Klemeš [1977a] pointed out, a comparison between two different alternatives is meaningful only if the operation and its cost were calculated with approximately the same accuracy for each alternative. A case in point is comparison of the operation of two reservoirs of different sizes, under consideration for construction at the same site. How should one discretize to obtain results of comparable accuracy? Klemeš studies this problem through computational experiments. His results indicate that it is the discretization interval (roughly capacity divided by the number of nodes) which must be kept constant, rather than the total number of grid nodes.

Our objective was to relate the error in the estimated control $|\Delta u|$ and cost of operation $|F' - F|$ to the discretization step ΔS . To obtain results of general applicability an analytical approach was followed. The basic assumptions in this analysis are that cost functions are smooth enough to satisfy some conditions of differentiability and that ΔS and Δu are small so that only leading terms need to be retained in power series expansions. In a strict mathematical sense, our results become exact asymptotically as $\Delta S \rightarrow 0$ but are quite accurate for finite values of ΔS . The results of such analyses have proven their usefulness in optimization [Luenberger, 1973] and numerical analysis [Dahlquist and Björck, 1974] because they yield results of practical usefulness even for finite values of the "small" parameter. For conventional DDP, the results are accurate if the stagewise cost and the cost-to-go functions may be approximated by quadratic functions in neighborhoods about the optimum and radii of the order of ΔS . For example, for the linear quadratic problem the analysis is exact for any value of ΔS .

In the case of conventional DDP, it has been shown that the error in control $|\Delta u|$ is of order ΔS , while the error in the calculated cost function $|F_{k+1}' - F_{k+1}|$ varies with ΔS^2 . Thus doubling the number of nodes should reduce Δu by a factor of about 2 and $|F_{k+1}' - F_{k+1}|$ by a factor of about 4. An interesting practical implication of $|\Delta u| \sim O(\Delta S)$ is in optimizing the short-term operation of a system where the (daily or weekly) release in each period is a small percentage of the reservoir capacity. The error in determining release through DDP may be large in relative terms, unless ΔS is much smaller than the typical value of volume released in a single stage. Another interesting implication is in comparing the operation of two reservoirs of different sizes. Assuming that the stagewise and cost-to-go functions are the same, our analysis indicates that comparable results are obtained when the same ΔS is used for both reservoirs, a conclusion which is in agreement with the observations of Klemeš [1977a].

The coefficient of proportionality in the linear relation between Δu and ΔS was shown to depend on the ratio of the second derivative of the cost-to-go function to the Hessian (sum of the second derivatives of the stagewise cost and cost-to-go functions). If the cost-to-go function is flat compared to the stagewise cost function, this coefficient is small and the results of DDP are accurate even for large ΔS . At the extreme (and usually unrealistic) case of practically linear cost-to-go function, DDP gives exact results even with the minimum possible number of nodes. For a highly curved cost-to-go function the same coarse grid would give results which are in serious error. Thus our analysis indicates that there is no single number of nodes which can be recommended for all cases. The appropriate number of nodes depends on the ratio of the curvatures (second derivatives) of the cost-to-go and the stagewise cost functions.

The calculated cost tends to the true minimum with quadratic order of convergence, since the error is of order ΔS^2 . Thus convergence in cost is always faster than convergence in control. This is fortunate for the practical usefulness of DDP, since it is the total operation cost we are concerned with. Quite often, seemingly serious deviations from the optimum control may produce not as serious a deterioration in the total operation cost. These features will be illustrated in the example for finite ΔS .

These results apply to deterministic as well as stochastic cases. An interesting implication of this analysis is that the effect of the state discretization error tends to be less in the stochastic case than in the deterministic case with the same economic parameters and mean inflow. This amelioration is the consequence of the smoothing associated with taking average values in the stochastic case.

4. GRADIENT DYNAMIC PROGRAMMING

In this section an apparently new computational technique for the solution of DDP problems is proposed. The primary motivation for its development was the desire to reduce the discretization error and thus achieve better accuracy than the one achieved by conventional DDP with the same grid. Another important objective was to take advantage of sophisticated nonlinear programming techniques to solve the single-stage optimization problem. While such techniques have found applications in successive-approximation versions of DP, they have been neglected in discrete DP.

Before examining a particular method, let us discuss the general issue of how a state discretization affects the calculated control and cost functions. As is illustrated by the analysis in section 2, the source of the discretization error can be traced back to the approximation of the cost to go (a function of continuous state variables) through interpolation from a finite number of grid points. In DDP, the value of the cost to go is calculated at each node and the cost-to-go function is approximated through piecewise linear interpolation. We have already shown that with this scheme the error in the calculated control is proportional to the discretization interval. Thus the key to improving the convergence rate would be to adopt a more accurate interpolation scheme.

The "most accurate or appropriate" interpolation scheme for a given problem depends on the smoothness properties of the function which is approximated and, in particular, its differentiability properties [see Davis, 1975]. For cost-to-go functions which hardly have continuous first derivatives, a piece-

wise linear approximation may be the wisest choice. However, for functions which have derivatives of high order (a "higher degree of smoothness" according to Davis [1975, p. 5]), more sophisticated interpolation schemes can reduce the approximation error quite significantly, and as we will illustrate with a particular case, reduce the error in the estimation of the control.

In GDP the cost to go and its derivatives with respect to state variables are calculated at all nodes. Then the cost-to-go function is piecewise approximated through the lowest-order polynomials which preserve at the surrounding nodes the calculated cost to go and its derivatives. The methodology is described for the multivariate case in the work by Foufoula-Georgiou and Kitanidis [1986]. In this paper we will first describe GDP in a simplified form applicable to the univariate case and then evaluate the effects of state discretization on control and cost.

With the terminology established in section 2, let

$$F_{k+1}(S_{i-1}), F_{k+1}(S_i), \left. \frac{dF_{k+1}}{dS} \right|_{S_{i-1}}$$

$$\left. \frac{dF_{k+1}}{dS} \right|_{S_i}$$

be the values and the first derivatives of the stage $(k + 1)$ cost-to-go function evaluated at the grid points S_{i-1} and S_i , where S_{i-1} and S_i define the interval within which the value $S(k) + q(k + 1) - u^*(k + 1)$ falls. Let G_{k+1} be the polynomial approximation of F_{k+1} in the interval $[S_{i-1}, S_i]$. The form of G_{k+1} will be determined so that

$$G_{k+1}(S_{i-1}) = F_{k+1}(S_{i-1}) \tag{21a}$$

$$G_{k+1}(S_i) = F_{k+1}(S_i) \tag{21b}$$

$$\left. \frac{dG_{k+1}}{dS} \right|_{S_{i-1}} = \left. \frac{dF_{k+1}}{dS} \right|_{S_{i-1}} \tag{21c}$$

$$\left. \frac{dG_{k+1}}{dS} \right|_{S_i} = \left. \frac{dF_{k+1}}{dS} \right|_{S_i} \tag{21d}$$

A computationally useful expression of the polynomial approximation $G_{k+1}(S)$ is of the form

$$G_{k+1}(S) = A(S)F_{k+1}(S_{i-1}) + B(S)F_{k+1}(S_i) + C(S) \left. \frac{dF_{k+1}}{dS} \right|_{S_{i-1}} + D(S) \left. \frac{dF_{k+1}}{dS} \right|_{S_i} \tag{22}$$

where $A(S)$ is 1 at S_{i-1} , 0 at S_i , and its first derivative is zero at S_{i-1} and S_i ; $B(S)$ is 0 at S_{i-1} , 1 at S_i , and its derivatives are zero at S_i and S_{i-1} ; $C(S)$ is 0 at S_{i-1} and S_i and its derivative $C'(S)$ is 1 at S_{i-1} and 0 at S_i ; and $D(S)$ is 0 at S_{i-1} and S_i , while $D'(S)$ is 0 at S_{i-1} and 1 at S_i . The lowest-order coefficient polynomials which satisfy these conditions are

$$A(S) = \frac{2}{(\Delta S)^3} \left[(S - S_{i-1}) + \frac{\Delta S}{2} \right] (S - S_i)^2 \tag{23a}$$

$$B(S) = \frac{2}{(\Delta S)^3} \left[(S_i - S) + \frac{\Delta S}{2} \right] (S - S_{i-1})^2 \tag{23b}$$

$$C(S) = \frac{1}{(\Delta S)^2} (S - S_{i-1})(S - S_i)^2 \tag{23c}$$

$$D(S) = \frac{1}{(\Delta S)^2} (S - S_i)(S - S_{i-1})^2 \tag{23d}$$

This method is known as Hermite interpolation (see, for example, Cheney [1982, p. 61]). It can be shown that of all the functions which reproduce the values of F_{k+1} and dF_{k+1}/dS at the nodes, the piecewise cubic approximation function defined by (22) and (23) is the one with the smallest average squared curvature [see Schultz, 1973, p. 31]. These relations can be generalized in the n -dimensional space in computational efficient ways [Kitanidis, 1986].

The first derivative of $G_{k+1}(S)$ is then expressed as

$$\frac{dG_{k+1}(S)}{dS} = \frac{dA}{dS} F_{k+1}(S_{i-1}) + \frac{dB}{dS} F_{k+1}(S_i) + \frac{dC}{dS} \frac{\partial F_{k+1}}{\partial S} \Big|_{S_{i-1}} + \frac{dD}{dS} \frac{\partial F_{k+1}}{\partial S} \Big|_{S_i} \quad (24)$$

where

$$\frac{dA}{dS} = \frac{6}{(\Delta S)^3} (S - S_i)(S - S_{i-1}) \quad (25a)$$

$$\frac{dB}{dS} = -\frac{6}{(\Delta S)^3} (S - S_i)(S - S_{i-1}) \quad (25b)$$

$$\frac{dC}{dS} = \frac{1}{(\Delta S)^2} \{(S - S_i)^2 + 2(S - S_{i-1})(S - S_i)\} \quad (25c)$$

$$\frac{dD}{dS} = \frac{1}{(\Delta S)^2} \{(S - S_{i-1})^2 + 2(S - S_{i-1})(S - S_i)\} \quad (25d)$$

Notice that $dA/dS + dB/dS = 0$ implying, as it was expected, that the solution depends on the difference (and not on the individual values) of the cost-to-go function at the end points of the interval. In contrast to conventional DDP, this approximation of the cost-to-go function has continuous first derivatives everywhere.

One of the major advantages of GDP is that it facilitates the application of the iterative Newton method for the solution of the single-stage nonlinear programming problem. Newton methods can be computationally very efficient. Their convergence rate improves with the accuracy of the available estimate of the second derivative of the objective function. Conventional DDP assumes piecewise linear cost-to-go function and, consequently, is of little use in calculating its second derivatives. In GDP d^2F_{k+1}/dS^2 is calculated from

$$\frac{d^2G_{k+1}(S)}{dS^2} = \frac{d^2A}{dS^2} F_{k+1}(S_{i-1}) + \frac{d^2B}{dS^2} F_{k+1}(S_i) + \frac{d^2C}{dS^2} \frac{\partial F_{k+1}}{\partial S} \Big|_{S_{i-1}} + \frac{d^2D}{dS^2} \frac{\partial F_{k+1}}{\partial S} \Big|_{S_i} \quad (26)$$

where

$$d^2A/dS^2 = \frac{12}{(\Delta S)^3} \left\{ (S - S_i) + \frac{\Delta S}{2} \right\} \quad (27a)$$

$$d^2B/dS^2 = -d^2A/dS^2 \quad (27b)$$

$$d^2C/dS^2 = \frac{2}{(\Delta S)^2} \{2(S - S_i) + (S - S_{i-1})\} \quad (27c)$$

$$d^2D/dS^2 = \frac{2}{(\Delta S)^2} \{2(S - S_{i-1}) + (S - S_i)\} \quad (27d)$$

Then, following a Newton iteration

$$[u^*(k+1)]^{i+1} = [u^*(k+1)]^i - \left[\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 G_{k+1}}{dS^2} \right]^{-1} \left[\frac{\partial c_k}{\partial u} - \frac{dG_{k+1}}{dS} \right] \quad (28)$$

where the first and second derivatives are calculated at values corresponding to $[u^*(k+1)]^i$. Iterations continue until convergence is achieved within the feasible region. If the search leads to an infeasible control, the following procedure is followed. The constraint which is violated becomes binding in which case it alone determines the optimum. Then, the Lagrange multiplier which corresponds to this constraint is calculated and if it is nonnegative the search is terminated (Kuhn-Tucker conditions are satisfied). Otherwise, the search is reinitialized starting from that point. The procedure converges under mild convexity requirements. (This procedure is generalized in the multidimensional case as the projected Newton method.) The Newton method has quadratic order of convergence, i.e., it has the asymptotically fastest rate of convergence among the commonly used gradient-based iterative methods. Here $u^*(k+1)$ is a function of $S(k)$. Then, the recursive equations which give $F_k[S(k)]$ and $dF_k[S(k)]/dS$ at the previous stage k are

$$F_k[S(k)] = c_k[S(k), u^*(k+1)] + G_{k+1}[S(k) + q(k+1) - u^*(k+1)] \quad (29)$$

$$\begin{aligned} \frac{dF_k}{dS} &= \frac{\partial c_k}{\partial S} + \frac{\partial c_k}{\partial u} \frac{du^*}{dS} + \frac{dG_{k+1}}{dS} \left(1 - \frac{du^*}{dS} \right) \\ &= \frac{\partial c_k}{\partial S} + \frac{dG_{k+1}}{dS} + \left(\frac{\partial c_k}{\partial u} - \frac{dG_{k+1}}{dS} \right) \frac{du^*}{dS} \end{aligned} \quad (30)$$

where the arguments in (30) have been omitted for notational simplicity. (For example, dF_k/dS is with respect to $S(k)$ and dG_{k+1}/dS is with respect to $S(k+1)$.) If no constraint is binding, $\partial c_k/\partial u - dG_{k+1}/dS = 0$ and (30) reduces to

$$dF_k/dS = \partial c_k/\partial S + (dG_{k+1})/dS \quad (31)$$

If a constraint is binding, $u^*(k+1)$ will be a (known) function of $S(k)$ denoted by $u^* = f[S(k)]$, and du^*/ds in (30) will be replaced by df/dS . For example, when one of the constraints (1b) becomes binding, (1a) results in a control policy $u^*(k+1)$ which is a linear function of $S(k)$, i.e., $df/dS = 1$, and therefore

$$dF_k/dS = \partial c_k/\partial S + \partial c_k/\partial u \quad (32)$$

5. PIECEWISE CUBIC APPROXIMATION OF THE CONTROL POLICY FOR GDP

Every discrete dynamic programming method estimates the control policy only at the grid points. For points in-between an approximation is needed. For the conventional DDP method a piecewise constant or a piecewise linear approximation is usually made. A more elaborate approximation, e.g., via a higher-order polynomial, would essentially require the estimation of the derivatives of the control policy at the grid points, a task which involves second and higher derivatives of the cost-to-go function. Since DDP assumes zero second derivatives of the cost-to-go function within intervals, such closer approximation of the control policy would be of limited value. This is not the case, however, for the proposed GDP method. In fact, cubic approximations of the control policy of GDP can be easily obtained. This requires the computation of

the first derivatives of the control policy on the grid points, and this computation is illustrated below.

If a constraint is binding, it determines u_{k+1}^* as well as du_{k+1}^*/dS . In the case of (1b),

$$(du_{k+1}^*)/dS = 1 \tag{33}$$

while in the case of (1c),

$$(du_{k+1}^*)/dS = 0 \tag{34}$$

If no constraint is binding u_{k+1}^* will be the solution of (4b). Taking partial derivatives of (4b) with respect to $S(k)$

$$\frac{\partial^2 c_k}{\partial S \partial u} + \frac{\partial^2 c_k}{\partial u^2} \frac{du_{k+1}^*}{\partial S} - \frac{d^2 F_{k+1}}{dS^2} \left(1 - \frac{du_{k+1}^*}{dS}\right) = 0 \tag{35}$$

from which

$$\frac{du_{k+1}^*}{dS} = \frac{-\frac{\partial^2 c_k}{\partial S \partial u} + \frac{d^2 F_{k+1}}{dS^2}}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \tag{36}$$

where the derivatives du_{k+1}^*/dS , $[\partial^2 c_k/\partial S \partial u]$, $\partial^2 c_k/\partial u^2$, and $d^2 F_{k+1}/dS^2$ are evaluated at $S(k)$, $[u^*(k+1)$ and $S(k)]$, $u^*(k+1)$, and $S(k) + q(k+1) - u^*(k+1)$, respectively. Naturally, in the above equation F_{k+1} is replaced by its approximation function G_{k+1} , i.e., a piecewise cubic function.

Having at every stage k the values u_{k+1}^* and the derivatives du_{k+1}^*/dS at all the points of the grid, a piecewise cubic approximation of the control policy can be obtained. For example, the approximation polynomial of the control in the interval $[S_{i-1}, S_i]$ would be

$$u_{k+1}^*(S) = A(S)u_{k+1}^*(S_{i-1}) + B(S)u_{k+1}^*(S_i) + C(S) \frac{du_{k+1}^*}{ds} \Big|_{S_{i-1}} + D(S) \frac{du_{k+1}^*}{\partial S} \Big|_{S_i} \tag{37}$$

where $A(S)$, $B(S)$, $C(S)$, and $D(S)$ are given, as before, from (23).

As it is shown theoretically in the next section (for ΔS asymptotically tending to zero) and as it is illustrated in the examples (for finite ΔS), GDP with only a small number of discrete states is as accurate in estimating the control policy as the conventional DDP with a much larger number of states. To benefit, however, from the coarser discretization that GDP permits, the approximation of the control function presented above becomes essential.

6. ERROR ANALYSIS OF GRADIENT DYNAMIC PROGRAMMING

Consider the true cost-to-go function $F_{k+1}(S)$ in the closed interval $[S_{i-1}, S_i]$ of length ΔS and its approximation (estimated cost-to-go function) $G_{k+1}(S)$ in the same interval. $F_{k+1}(S)$ and $G_{k+1}(S)$ satisfy conditions (21). Furthermore, for the purposes of the error analysis, it is assumed that $F_{k+1}(S)$ has continuous fourth derivatives. Before proceeding with the error analysis, the following lemma is needed.

Lemma. If $G_{k+1}(S)$ is a polynomial approximation of $F_{k+1}(S)$ in an interval of length (ΔS) , and $G_{k+1}(S)$ is so defined as to preserve the values $F_{k+1}(S)$ and the values of the first derivatives $dF_{k+1}(S)/dS$ at the end points (S_{i-1} and S_i) of the interval (condition (21)), then for any point within the interval

$$|F_{k+1} - G_{k+1}| \leq \frac{1}{384} \left| \frac{d^4 F_{k+1}}{dS^4} \right| (\Delta S)^4 \tag{38}$$

$$\left| \frac{dF_{k+1}}{dS} - \frac{dG_{k+1}}{dS} \right| \leq \frac{1}{72\sqrt{3}} \left| \frac{d^4 F_{k+1}}{dS^4} \right| (\Delta S)^3 \tag{39}$$

where $d^4 F_{k+1}/dS^4$ is evaluated at a point within the interval. The proof of the above lemma is given in the appendix. Equations (38) and (39) will now be used in the error analysis of the gradient dynamic programming method.

Using the established terminology and assuming that F_{k+1} and its derivative are known without error at the grid nodes one can write

$$\frac{\partial c_k}{\partial u} \Big|_{u^*(k+1)} - \frac{dG_{k+1}}{dS} \Big|_{S(k)+q(k+1)-u^*(k+1)} = 0 \tag{40}$$

$$\frac{\partial c_k}{\partial u} \Big|_{U^*(k+1)} - \frac{dF_{k+1}}{dS} \Big|_{S(k)+q(k+1)-U^*(k+1)} = 0 \tag{41}$$

Expanding the terms of (40) around the true optimal policy $U^*(k+1)$ and keeping only the lowest-order terms

$$\frac{\partial c_k}{\partial u} \Big|_{U^*(k+1)} + \frac{\partial^2 c_k}{\partial u^2} \Big|_{U^*(k+1)} \Delta u_{k+1} - \frac{dG_{k+1}}{dS} \Big|_{A^* \equiv S(k)+q(k+1)-U^*(k+1)} + \frac{d^2 G_{k+1}}{dS^2} \Big|_{A^*} \Delta u_{k+1} = 0 \tag{42}$$

Subtracting (41) from (42)

$$\Delta u_{k+1} \left(\frac{\partial^2 c_k}{\partial u^2} \Big|_{U^*(k+1)} + \frac{d^2 G_{k+1}}{dS^2} \Big|_{A^*} \right) = \left(\frac{dG_{k+1}}{dS} \Big|_{A^*} - \frac{dF_{k+1}}{dS} \Big|_{A^*} \right) \Delta u_{k+1} \tag{43}$$

Then using (39), neglecting the higher-order term $\Delta u \Delta S^2$ and taking absolute values,

$$|\Delta u_{k+1}| \leq \frac{1}{72\sqrt{3}} \left| \frac{(d^4 F_{k+1})/dS^4}{\frac{\partial^2 c_k}{\partial u^2} + \frac{d^2 F_{k+1}}{dS^2}} \right| \Delta S^3 \equiv |\Delta u_{k+1}|^{\max}(\text{GDP}) \tag{44}$$

where the derivatives $\partial^2 c_k/\partial u^2$, $d^2 G_{k+1}/dS^2$, and $\partial^4 F_{k+1}/\partial S^4$ are evaluated at $U^*(k+1)$, $S(k) + q(k+1) - U^*(k+1)$, and a point in the interval within which $S(k) + q(k+1) - u^*(k+1)$ falls, respectively. Thus $\Delta u_{k+1} \sim O(\Delta S^3)$.

The cost to go may now be calculated from (29) rewritten as

$$F_k'[S(k)] = c_k[S(k), u^*(k+1)] + G_{k+1}[S(k) + q(k+1) - u^*(k+1)] \tag{45}$$

while the truly optimum satisfies (13). Subtracting (13) from (45)

$$\begin{aligned} F_k'[S(k)] - F_k[S(k)] &= c_k[S(k), u^*(k+1)] - c_k[S(k), U^*(k+1)] \\ &+ G_{k+1}(A^* - \Delta u_{k+1}) - F_{k+1}(A^*) \\ &= \frac{\partial c_k}{\partial u} \Big|_{U^*} \Delta u_{k+1} + G_{k+1}(A^*) - \frac{dG_{k+1}}{dS} \Big|_{A^*} \Delta u_{k+1} - F_{k+1}(A^*) \end{aligned} \tag{46a}$$

where terms of order higher than ΔS^4 have been neglected. Using (4b) and keeping again the lowest-order terms

$$F_k'[S(k)] - F_k[S(k)] = G_{k+1}(A^*) - F_{k+1}(A^*) \quad (46b)$$

and from (38)

$$|F_k'[S(k)] - F_k[S(k)]| \leq \frac{1}{384} \left| \frac{d^4 F_{k+1}}{dS^4} \right| \Delta S^4 \quad (47)$$

The derivative of the cost to go, for no constraint binding, satisfies (31) rewritten as

$$\frac{dF_k'}{dS} = \frac{\partial c_k}{\partial S} \Big|_{u^*} + \frac{dG_{k+1}}{dS} \Big|_{S(k)+q(k+1)-u^*(k+1)} \quad (48)$$

while the actual satisfies

$$\frac{dF_k}{dS} = \frac{\partial c_k}{\partial S} \Big|_{u^*} + \frac{dF_{k+1}}{dS} \Big|_{A^*} \quad (49)$$

Subtracting (49) from (48), expanding and keeping the lowest-order terms

$$\begin{aligned} \frac{dF_k'}{dS} - \frac{dF_k}{dS} &= \frac{\partial^2 c_k}{\partial u \partial S} \Big|_{u^*} \Delta u_{k+1} + \frac{dG_{k+1}}{dS} \Big|_{A^*} \\ &\quad - \frac{d^2 G_{k+1}}{dS^2} \Big|_{A^*} \Delta u_{k+1} - \frac{dF_{k+1}}{dS} \Big|_{A^*} \\ &= \left(\frac{\partial^2 c_k}{\partial u \partial S} \Big|_{u^*} - \frac{d^2 F_{k+1}}{dS^2} \right) \Delta u_{k+1} \\ &\quad + \frac{dG_{k+1}}{dS} \Big|_{A^*} - \frac{dF_{k+1}}{dS} \Big|_{A^*} \sim O(\Delta S^3) \end{aligned} \quad (50)$$

where we used that

$$\frac{d^2 G_{k+1}}{dS^2} \Delta u_{k+1} = \frac{d^2 F_{k+1}}{dS^2} \Delta u_{k+1} + O(\Delta S^5)$$

In the derivation of (44), (47), and (50) F_{k+1} and dF_{k+1}/dS were assumed known without error. This should be the case at the last stage, i.e., for $k + 1 = N$. At other steps, the approximately calculated nodal values F_{k+1}' and dF_{k+1}'/dS are used in the place of F_{k+1} and dF_{k+1}/dS in (22). However, since $|F_{k+1}' - F_{k+1}| \sim O(\Delta S^4)$ and $|dF_{k+1}'/dS - dF_{k+1}/dS| \sim O(\Delta S^3)$, one may easily verify by repeating the analysis of (40) through (50) that

$$|\Delta u_{k+1}| \sim O(\Delta S^3) \quad (51a)$$

$$|F_k' - F_k| \sim O(\Delta S^4) \quad (51b)$$

$$\left| \frac{dF_k'}{dS} - \frac{dF_k}{dS} \right| \sim O(\Delta S^3) \quad (51c)$$

Thus while the convergence of DDP is linear, the convergence of GDP is of order three. However, the reader is reminded that the improved order of convergence of the GDP is predicated on a higher degree of smoothness of the cost-to-go function than that required by DDP. Thus in the expansions into Taylor series for GDP we assumed that the cost to go is four times continuously differentiable, while for DDP that it is twice continuously differentiable.

7. EXAMPLE OF A SINGLE-RESERVOIR OPERATION SYSTEM

The system under study is a simplified representation of the Saylorville Reservoir, on the Des Moines River, Iowa. The dam is located 214 miles upstream from the mouth of the Des Moines River on the Mississippi River and about 9 miles

upstream from the city of Des Moines, Iowa. At full flood control pool, elevation 890 ft above mean sea level (msl), the lake extends 54 miles upstream from the dam, occupies about 16,700 acres of land, and the storage is 670,000 acft. At conservation pool, elevation 833 ft above msl, the lake extends for about 17 miles upstream, occupies 5400 acres, and corresponds to storage 74,000 ac-ft. The project is owned by the federal government and is operated and maintained by the U.S. Army Corps of Engineers, Rock Island District.

The optimization problem studied here is a finite-horizon, short-term optimal control problem. The operating horizon consists of five periods, each period equal to a fortnight. The control variables are the amounts of water released from each reservoir and the state variables are the amounts of water stored in each reservoir. All variables are expressed in units of 1000 ac-ft. The operating cost consists of a terminal cost c_T and stagewise costs c_k . The role of the terminal cost is to make the short-term operation consistent with the long-term objectives. It is assumed that c_T is a linear function of the amount of water remaining in the reservoir at the end of the operating horizon,

$$c_T = aS(T) \quad (52)$$

where a is a cost coefficient assumed equal to 1/150. The stagewise cost c_k , which represents flood damages, is assumed a function of the release $u(k + 1)$ at stage k . Approximately,

$$\begin{aligned} c_k &= 0 & u(k + 1) &\leq R \\ c_k &= \left(\frac{u(k + 1) - R}{R} \right)^3 & u(k + 1) &> R \end{aligned} \quad (53)$$

$$k = 0, 1, 2, 3, 4$$

where R is the maximum "no damage" release taken equal to 140. The capacity of the reservoir is $K = 600$. The system dynamics and constraints are described in (1a)–(1c) for $k = 0, 1, 2, 3, 4$.

The performance criterion is

$$J = \sum_{k=0}^4 c_k[u(k + 1)] + aS(5) \quad (54)$$

and releases are determined by minimizing the value of J (or the expected value of the cost to go at each stage, for the stochastic case).

The above optimal control problem has been solved by both the conventional DDP and the proposed GDP methods. The purpose was to compare the estimated control policies for several state discretization schemes to the "true" control policy of the system. The true control policy was computed numerically using a very fine discretization scheme (number of states, $NS = 62$) and was for all practical purposes the same for both DDP and GDP methods. The true optimum was also checked with nonlinear programming. The discretization scheme of Savarenkyi (see, for example, *Klemeš [1977b]*) was used. With this scheme, the storage values of zero (empty reservoir) and K (full reservoir) are considered separate states.

Two examples, one with deterministic input and the other with stochastic input of the same mean and a specified lognormal marginal probability distribution, have been studied. The inflows to the system have been assumed statistically independent. This assumption, although not realistic for biweekly flows, serves well the illustrative purpose of our case studies.

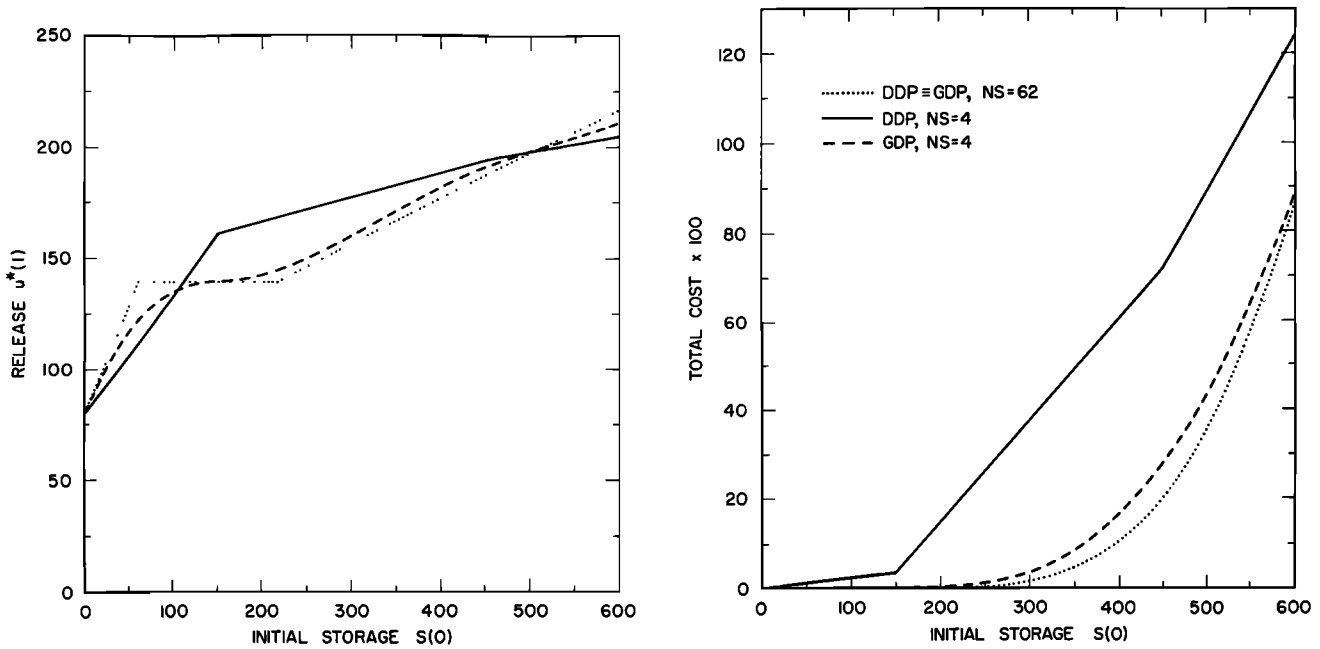


Fig. 3. Optimal release $u^*(1)$ and total cost as functions of the initial storage $S(0)$. Deterministic case; state discretization scheme of four nodes.

7.1. Deterministic Case

The input to the system over the five operating periods is assumed to have the shape of a symmetric hydrograph with values

$$q(1) = 80 \quad q(2) = 100 \quad q(3) = 130$$

$$q(4) = 100 \quad q(5) = 80$$

In all cases we enforced the constraint that the release should not be less than the smaller of R and available water, $S(k) + q(k + 1)$. Figures 3, 4, and 5 show a comparison of the optimal release $u^*(1)$ during the first operating period, as a function of the initial storage $S(0)$ for three state discretization schemes with 4, 8, and 14 discrete nodes ($NS = 4, 8, \text{ and } 14$), respectively.

It is observed that the conventional DDP with $NS = 4$ yields an optimal control policy which considerably differs from the true one. As NS increases, the maximum error decreases and maximum $|\Delta u(\text{DDP})| \sim 1/NS$ (linear rate of convergence). GDP yields a policy which is very close to the true one, even for NS as small as 4. Furthermore, the improvement from $NS = 4$ to $NS = 8$ is definitely superlinear. Figures 3, 4, and 5 compare the total costs for the same discretization configurations, as calculated by each methodology. It is obvious that GDP is near optimal even with as few as four nodes. This is particularly interesting since the cost-to-go function does not have continuous higher-order derivatives in this optimization problem.

7.2. Stochastic Case

The previous example was further solved for stochastic inflows, having at each operating period i a lognormal distribution with mean $\bar{q}(i)$ and variance $\text{Var}(q(i))$, where

$$\bar{q}(i) = (80, 100, 130, 100, 80)$$

$$\text{Var}(q(i)) = (900, 900, 900, 900, 900)$$

The probability distribution of the input was represented with 10 discrete values placed at equal probability intervals apart. Since our objective is to study the effect of state discretization, we will assume that this representation is adequate for our purposes. The acceptable probability of violating the nonnegativity or capacity constraints were set equal to five percent. Then, the objective is to minimize

$$F_k[S(k)] = \min_{u(k+1)} \{c_k[u(k+1)] + \sum_{i=1}^l p_i F_{k+1}[S(k) + q_i(k+1) - u(k+1)]\} \quad (55)$$

as compared to the deterministic case, in which

$$F_k = \min_{u(k+1)} \{c_k[u(k+1)] + F_{k+1}[S(k) + \bar{q}(k+1) - u(k+1)]\} \quad (56)$$

The weights $p_i, i = 1, 1, \dots, l$ in (55) correspond to the probabilities of having inflow within the i th interval of the discrete representation of inflows. Thus the summation in (55) extends over values of F_{k+1} computed at points around the value $S(k) + \bar{q}(k+1) - u(k+1)$.

Figures 6 and 7 compare the optimal release $u^*(1)$ obtained for two discretization schemes with four and eight states, respectively, to the true optimal release. The corresponding calculated mean operating costs are also compared in the same figures. It is obvious that GDP yields more accurate solutions than DDP. By comparing the accuracy of calculated release policies in Figures 6 and 7 one may confirm the linear convergence of DDP and the superlinear convergence of GDP.

By comparing Figures 3 and 7 it becomes obvious that for the same discretization levels both DDP and GDP are more accurate in the stochastic than in the deterministic case. This should be attributed to the smoothing effect of averaging (equation (55)). For example, in DDP, the weighted second

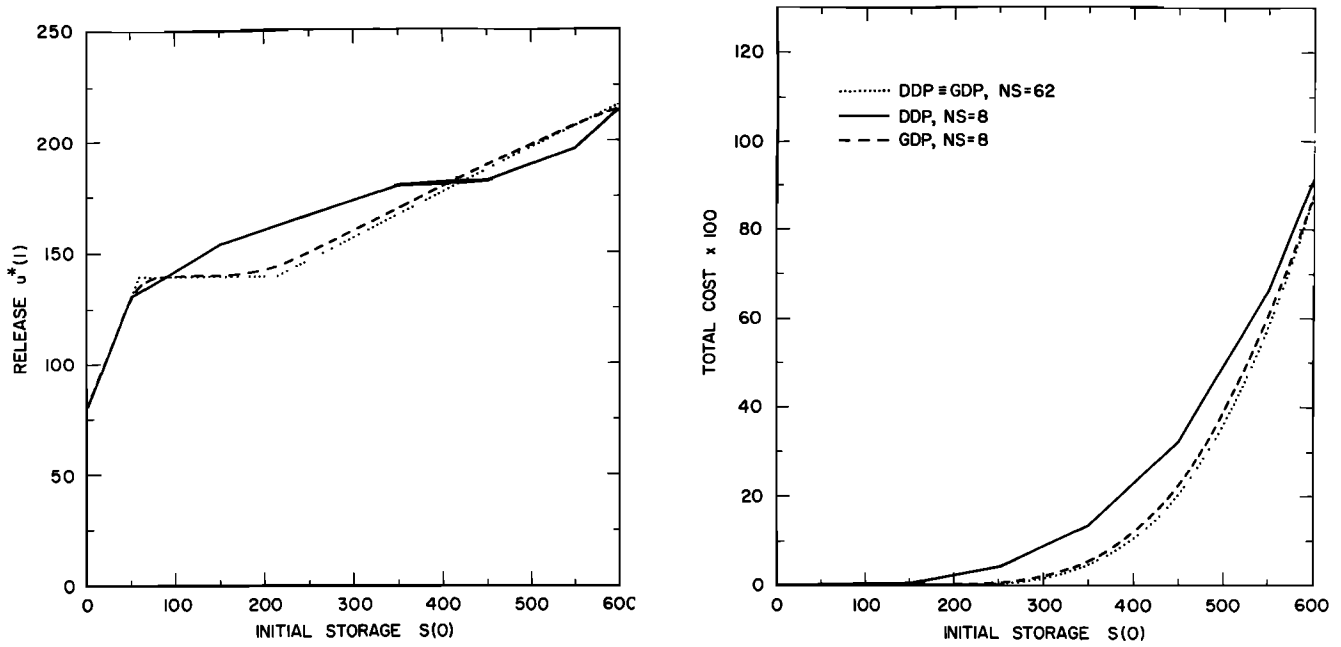


Fig. 4. Optimal release $u^*(1)$ and total cost as functions of the initial storage $S(0)$. Deterministic case; state discretization scheme of eight nodes.

derivative

$$\sum_{i=1}^I p_i \left. \frac{d^2 F_{k+1}}{dS^2} \right|_{S(k)+q_i(k+1)-U^*(k+1)}$$

takes less extreme values than

$$\left. \frac{d^2 F_{k+1}}{dS^2} \right|_{S(k)+\bar{q}(k+1)-U^*(k+1)}$$

and, consequently, the discretization error should be smaller in the stochastic case.

In stochastic DDP, the linear nonnegativity and capacity constraints are replaced by their deterministic equivalents [see *Stedinger et al.* [1984] and their list of references). Therefore when a constraint is binding the stochastic optimal control differs from the deterministic one by an amount depending on the allowable probability of violating the constraints. This probability was taken equal to 5%. In the example considered, the nonnegativity constraint becomes binding much more often than the capacity constraint. Therefore for small initial storages, the stochastic case results in lower (more conserva-

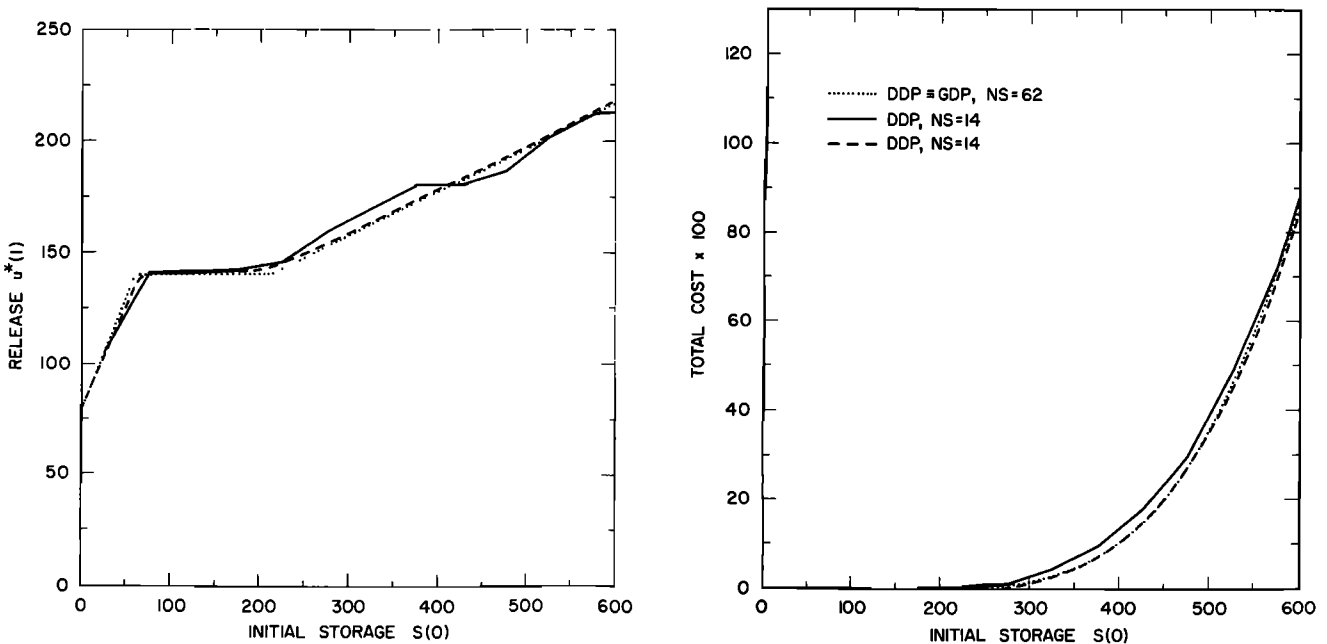


Fig. 5. Optimal release $u^*(1)$ and total cost as functions of the initial storage $S(0)$. Deterministic case; state discretization scheme of fourteen nodes.

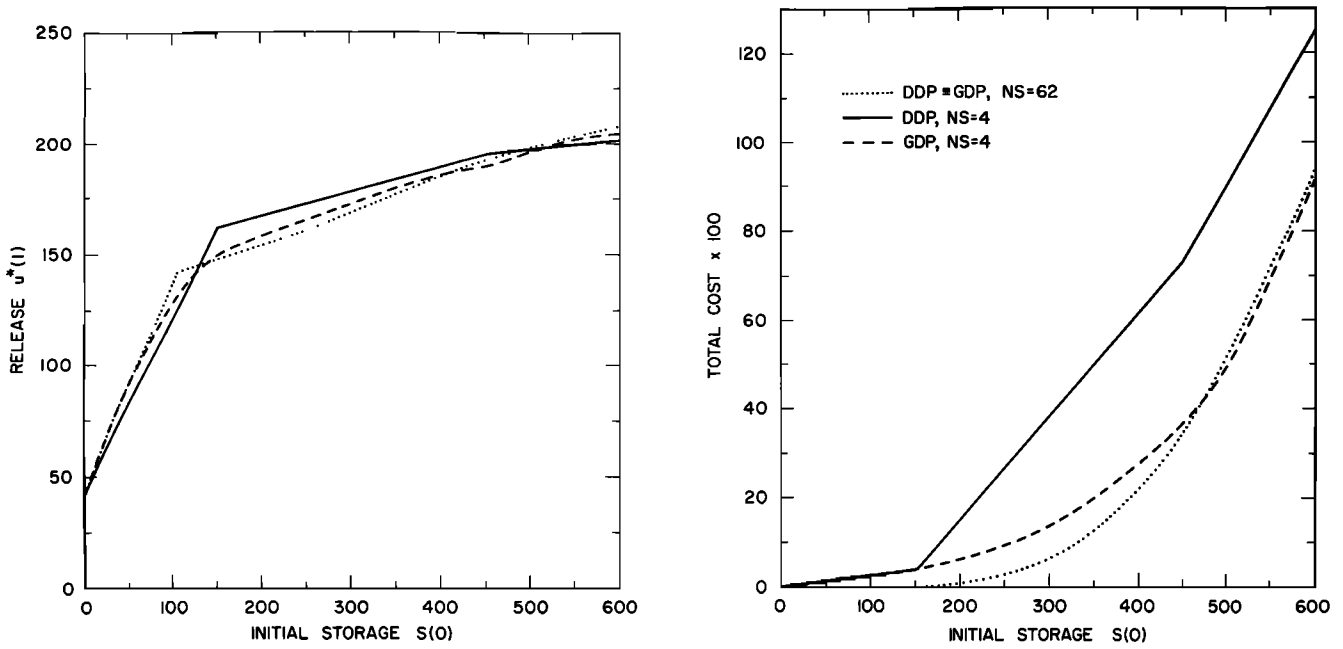


Fig. 6. Optimal release $u^*(1)$ and total cost as functions of the initial storage $S(0)$. Stochastic case; state discretization scheme of four nodes.

tive) releases, and consequently higher costs, as compared to the analogous deterministic case. All of the above points can be clearly observed in Figures 6 and 7.

8. SUMMARY AND CONCLUSIONS

The effect of state discretization on estimated control variables (such as releases) and cost functions was studied for two versions of discrete dynamic programming. Analytical expressions were obtained assuming small discretization interval and the results were verified in a case study involving the

optimization of reservoir operation.

First, the effect of state discretization in conventional DDP was studied. In DDP the cost to go is calculated at each node of the discretization grid and is approximated through linear interpolation for points between nodes. The error in the control was shown to vary linearly with the discretization interval and the error in the cost function to vary with the square of the discretization interval. The coefficients of proportionality depend on the ratio of the second derivatives ("curvature") of the cost-to-go functions over the second derivative of the

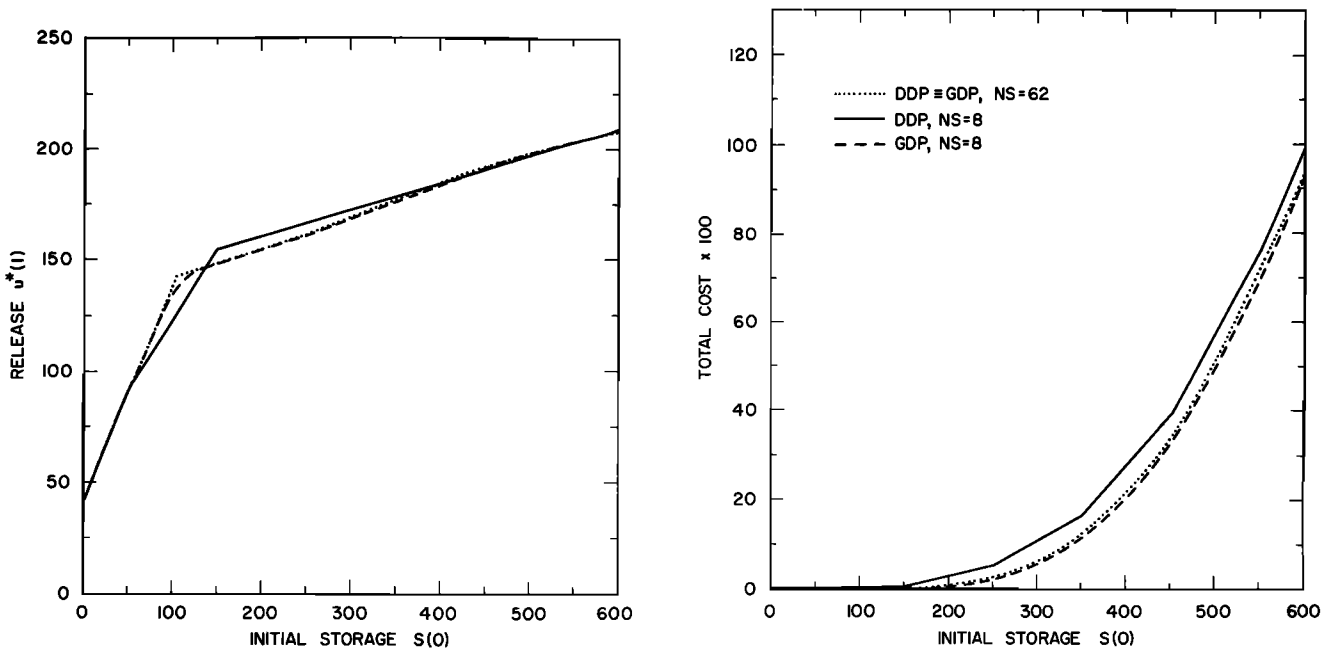


Fig. 7. Optimal release $u^*(1)$ and total cost as functions of the initial storage $S(0)$. Stochastic case; state discretization scheme of eight nodes.

stagewise cost. The analysis thus has shown that there is no single number of nodes which would achieve the same accuracy for all cases. On the basis of bounds which are analytically derived, one can determine whether results obtained from different cases are of comparable accuracy. In the stochastic case it is the expected value of second derivative of the cost to go which is relevant and the effect of state discretization tends to be less in the stochastic case than in the deterministic one.

Gradient dynamic programming is discrete dynamic programming in which the cost-to-go function F and its first derivatives with respect to the state variables are calculated at all nodes. Between nodes, F is approximated with the lowest-order polynomial which preserves both the values of F and its first derivatives calculated at the nodes. For smoothly varying functions this approximation is much more accurate than the piecewise linear approximation of conventional DDP. The error in the control was shown to vary with the third power of the discretization interval and the error in the cost to vary with the fourth power of the discretization interval.

APPENDIX: PROOF OF LEMMA (EQUATIONS (38) AND (39)) IN SECTION 5

Consider a function $h(x)$, $x \in [x_1, x_2]$, which has continuous fourth derivatives and at the end points of the interval $[x_1, x_2]$ satisfies the conditions

$$\begin{aligned} h(x_1) &= h(x_2) = 0 \\ h'(x_1) &= h'(x_2) = 0 \end{aligned} \quad (\text{A1})$$

where h' denotes the derivative of h . It will be shown that $h(x)$ is of the order Δx^4 , where $\Delta x = x_2 - x_1$.

Performing a Taylor series expansion about x_1 and using conditions (A1)

$$h(x) = \frac{1}{2}h''(x_1)(x - x_1)^2 + \frac{1}{6}h^{(3)}(x_1)(x - x_1)^3 + \frac{1}{24}h^{(4)}(\zeta)(x - x_1)^4 \quad (\text{A2})$$

$$h'(x) = h''(x_1)(x - x_1) + \frac{1}{2}h^{(3)}(x_1)(x - x_1)^2 + \frac{1}{6}h^{(4)}(\zeta)(x - x_1)^3 \quad (\text{A3})$$

where ζ lies between x_1 and x . Setting $x = x_2$ and using conditions (A1)

$$0 = \frac{1}{2}h''(x_1)\Delta x^2 + \frac{1}{6}h^{(3)}(x_1)\Delta x^3 + \frac{1}{24}h^{(4)}(\zeta_2)\Delta x^4 \quad (\text{A4})$$

$$0 = h''(x_1)\Delta x + \frac{1}{2}h^{(3)}(x_1)\Delta x^2 + \frac{1}{6}h^{(4)}(\zeta_2)\Delta x^3 \quad (\text{A5})$$

where ζ_2 lies between x_1 and x_2 . Solving (A4) and (A5) in terms of $h''(x_1)$ and $h^{(3)}(x_1)$

$$h''(x_1) = \frac{1}{12}h^{(4)}(\zeta_2)\Delta x^2 \quad (\text{A6})$$

$$h^{(3)}(x_1) = -\frac{1}{2}h^{(4)}(\zeta_2)\Delta x \quad (\text{A7})$$

Substituting in (A2) and (A3)

$$h(x) = \frac{1}{24}[\Delta x - 2(x - x_1)](x - x_1)^2\Delta x h^{(4)}(\zeta_2) + \frac{1}{24}(x - x_1)^4 h^{(4)}(\zeta) \quad (\text{A8})$$

$$h'(x) = \frac{1}{12}[\Delta x - 3(x - x_1)](x - x_1)\Delta x h^{(4)}(\zeta_2) + \frac{1}{6}(x - x_1)^3 h^{(4)}(\zeta) \quad (\text{A9})$$

It is obvious that if the fourth derivative of h is zero, h vanishes everywhere in the interval $[x_1, x_2]$. Otherwise, $h(x) \sim O(\Delta x^4)$ and $h'(x) \sim O(\Delta x^3)$. To obtain analytical ex-

pressions of the bounds, consider the case of small Δx . Because of the continuity of the fourth derivative, $h^{(4)}(\zeta_2)$ tends to $h^{(4)}(\zeta)$ as Δx decreases and ζ_2 tends to ζ . Consequently,

$$h(x) = \frac{1}{24}[\Delta x^2(x - x_1)^2 - 2\Delta x(x - x_1)^3 + (x - x_1)^4]h^{(4)}(\zeta) \quad (\text{A10})$$

$$h'(x) = \frac{1}{12}[\Delta x^2(x - x_1) - 3\Delta x(x - x_1)^2 + 2(x - x_1)^3]h^{(4)}(\zeta) \quad (\text{A11})$$

Maximizing the absolute values of these expressions

$$|h(x)| \leq \frac{1}{384}h^{(4)}(\zeta)\Delta x^4 \quad (\text{A12})$$

$$|h'(x)| \leq \frac{1}{72\sqrt{3}}h^{(4)}(\zeta)\Delta x^3 \quad (\text{A13})$$

Then, by defining $h(x) \equiv F_{k+1}(x) - G_{k+1}(x)$, the difference between the actual cost to go and its approximation, we obtain inequalities (38) and (39).

Acknowledgments. The material is based upon work supported by the National Science Foundation under grants CEE-8212066 and CEE-8420664. Comments by J. Stedinger and anonymous reviewers to an earlier version of the paper are appreciated.

REFERENCES

- Askew, A. J., Optimum reservoir operating policies and the imposition of a reliability constraint, *Water Resour. Res.*, 10(1), 51-56, 1974.
- Bellman, R. E., On the theory of dynamic programming, *Proc. Nat. Acad. Sci.*, 38, 716-719, 1952.
- Bellman, R. E., *Dynamic Programming*, Princeton University Press, Princeton, N. J., 1957.
- Bellman, R. E., and S. E. Dreyfus, *Applied Dynamic Programming*, Princeton University Press, Princeton, N. J., 1962.
- Buras, N., Conjunctive operation of dams and aquifers, *J. Hydraul. Eng.*, 89(HY6), 111-131, 1963.
- Butcher, W., Stochastic dynamic programming for optimum reservoir operation, *Water Resour. Bull.*, 17(1), 115-123, 1971.
- Cheney, E. W., *Approximation Theory*, Chelsea Publishing, New York, 1982.
- Collado, J. M., Parameter and forecasting uncertainties in the optimization of reservoir operation, Ph.D. thesis, Dep. of Civ. and Environ. Eng., Univ. of Iowa, Iowa City, September 1984.
- Dahlquist, G., and A. Bjorck, *Numerical Methods*, translated by N. Anderson, Prentice-Hall, Englewood Cliffs, N. J., 1974.
- Davis, P. J., *Interpolation and Approximation*, 393 pp., Dover, New York, 1975.
- Doran, D. G., An efficient transition definition for discrete state reservoir analysis: The divided interval technique, *Water Resour. Res.*, 11(6), 867-873, 1975.
- Foufoula-Georgiou, E., and P. Kitanidis, Constrained gradient dynamic programming for stochastic optimal control of water resources systems, technical note, St. Anthony Falls Hydraul. Lab., Univ. of Minn., Minneapolis, 1986.
- Gablinger, M., and D. P. Loucks, Markov models for flow regulation, *J. Hydraul. Eng.*, 96(HY1), 165-181, 1970.
- Kitanidis, P. K., Real-time forecasting of river flows and stochastic optimal control of multireservoir systems, *Rep. 258*, Iowa Inst. of Hydraul. Res., Univ. of Iowa, Iowa City, 1983.
- Kitanidis, P. K., A first-order approximation to closed-loop stochastic optimal control of reservoirs, technical note, St. Anthony Falls Hydraul. Lab., Univ. of Minn., Minneapolis, 1985.
- Kitanidis, P. K., Hermite interpolation with a rectangular grid in n -dimensional space, technical note, St. Anthony Falls Hydraul. Lab., Univ. of Minn., Minneapolis, 1986.
- Klemeš, V., Discrete representation of storage for stochastic reservoir optimization, *Water Resour. Res.*, 13(1), 149-158, 1977a.
- Klemeš, V., Value of information in reservoir optimization, *Water Resour. Res.*, 13(5), 837-850, 1977b.
- Loaiciga, H. A., and M. A. Mariño, An approach to parameter estimation and stochastic control in water resources with an appli-

- cation to reservoir operation, *Water Resour. Res.*, 21(11), 1575-1585, 1985.
- Loucks, D. P., and L. M. Falkson, A comparison of some dynamic, linear, and policy iteration methods for reservoir operation, *Water Resour. Bull.*, 6(3), 384-400, 1970.
- Luenberger, D. G., *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass., 1973.
- Massé, P., *Les Réserves et a Regulation de l'Avenir Dans la Vie Economique*, Herman, Paris, 1946.
- Moran, P. A. P., *The Theory of Storage*, Methuen, London, 1959.
- Savarenskiy, A. D., Metod rascheta regulirovaniya stoka, *Gidrotekh. Stroit.*, 2, 24-28, 1940.
- Schultz, M. H., *Spline Analysis*, 156 pp., Prentice-Hall, Englewood Cliff, N. J., 1973.
- Stedinger, J. R., B. F. Sula, and D. P. Loucks, Stochastic dynamic programming models for reservoir operation optimization, *Water Resour. Res.*, 20(11), 1499-1505, 1984.
- Su, S., and R. Deininger, Modeling the regulation of Lake Superior under uncertainty of future water supplies, *Water Resour. Res.*, 10(1), 11-25, 1974.
- Wasimi, S. A., and P. K. Kitanidis, Real-time forecasting and daily operation of a multireservoir system during floods by linear quadratic Gaussian control, *Water Resour. Res.*, 19(6), 1511-1522, 1983.
- Yakowitz, S., Dynamic programming applications in water resources, *Water Resour. Res.*, 18(4), 673-696, 1982.
- Yeh, W., Reservoir management and operation models: A state-of-the-art review, *Water Resour. Res.*, 21(2), 1797-1818, 1985.
- Zaphirakos, A. N., Operation of a multipurpose reservoir under drought conditions, M.S. thesis, Dep. of Civ. and Environ. Eng., Univ. of Iowa, Iowa City, July 1982.
-
- E. Foufoula-Georgiou, Department of Civil Engineering, Iowa State University, Ames, IA 50011.
- P. K. Kitanidis, St. Anthony Falls Hydraulic Laboratory, Department of Civil and Mineral Engineering, Mississippi River at Third Avenue, S.E., Minneapolis, MN 55414.

(Received November 3, 1986;
accepted November 11, 1986.)