Stochastic order-based optimal design of a surface reservoir–irrigation district system

Hosein Alizadeh and S. Jamshid Mousavi

ABSTRACT

This paper explores the effects of streamflow uncertainty and the type of stochastic order, which is used for comparing stochastic variables, on optimal design of a reservoir multi-crop irrigation district system. Four nonlinear mathematical programs with an economic objective function including deterministic, stochastic-EXP with expected value order, stochastic-SD with stochastic dominance (SD) order, and stochastic-EGCL with expected gain-confidence limit (EGCL) order were developed. Afterwards, the approaches of successive linear programming (SLP) and PSO–MC, which combines particle swarm optimization (PSO) algorithm and Monte Carlo simulation (MC), to solve the programs were selected. Hajiarab Irrigation District located in Ghazvin Province of Iran was used as a case study and the results obtained from using different programs and solution approaches were analyzed and compared. Among the solution approaches, SLP could not solve the programs other than the deterministic one while PSO–MC could find at least good solutions, if not the global optima, for all the programs. Among the main decision variables, the stochastic programs resulted in a reduced size of irrigation district compared with that obtained by the deterministic program. Moreover, the programs using stochastic orders other than a simple expected value converged at solutions different from the solution reached by the expected value-based program.

Key words | economic analysis, irrigation planning, Monte Carlo simulation, particle swarm optimization, reservoir design, stochastic orders

ABBREVIATIONS

B/C Benefit-over-cost ratio
Deterministic Deterministic mathematical program
DP Dynamic programming
EGCL Expected gain-confidence limit order
FSD First-degree stochastic dominance
GSP General stochastic program
IR Iranian Rial monetary unit
IRR Internal rate of return
LP Linear programming
MC Monte Carlo simulation
NLP Nonlinear programming
NPV Present value of net benefit or net present value
O&M Operation and maintenance
PDF Probability density function
CDF Cumulative distribution function
PSO Particle swarm optimization
PSO A solution approach consisting of two modules of PSO-MC and MC
r.h.s. Right hand side
SD Stochastic dominance
SDP Stochastic dynamic programming
SDP-DP A model combining SDP and DP sub-models
SDP-LP A model combining SDP and LP sub-models
SLP Successive linear programming
SP Stochastic programming
**SSD** Second-degree stochastic dominance

**Stochastic-EXP** Stochastic mathematical program using expected value order

**Stochastic-SD** Stochastic mathematical program using SD order

**Stochastic-EGCL** Stochastic mathematical program using EGCL order

**TSD** Third-degree stochastic dominance

**NOTATION**

- \( BF_t \): Baseline flow required to support the ecosystem water requirement at downstream of the reservoir in month \( t \)
- \( CP_{c,t} \): Net irrigation water requirement at farm per unit area of crop \( c \) in month \( t \)
- \( CP_c \): Fraction representing the share of crop \( c \) in total cultivation area
- \( CPC_c \): Production cost of crop \( c \) per unit area
- \( E(.) \): Expected value function
- \( F_i(.) \): Cumulative distribution function of economic return of project \( i \)
- \( IA_{x,y} \): Total irrigated area in the year \( y \) of the scenario (outcome) \( s \)
- \( IE \): Irrigation efficiency factor
- \( K \): Reservoir capacity
- \( LB \): Lower bound for crop pattern
- \( MaxLand \): The maximum irrigable land for cultivation
- \( NPV \): Present value of net benefit achieved from irrigation over a long-term horizon
- \( Pe_c \): Price of crop \( c \) per unit mass
- \( PVB \): Present value of benefit
- \( PVC \): Present value of cost
- \( Q_{x,y,t} \): Volume of water entered into the reservoir in month \( t \) of year \( y \) of scenario \( s \)
- \( R_{x,y,t} \): Volume of water released from the reservoir in month \( t \) of year \( y \) of scenario \( s \)
- \( S_{x,y,t} \): Reservoir storage at the beginning of month \( t \)
- \( TIA \): Area of the irrigation district
- \( UB \): Upper bound for crop pattern
- \( Y_c \): Mass of actual yield of crop \( c \) per unit cultivated area

\( \Omega \): Represents a stochastic order

\( a_1 \): A parameter associated with the variable cost term for dam construction

\( a_2 \): A parameter associated with the construction cost per unit area of the irrigation district

\( b_1 \): A parameter associated with the fixed cost term for dam construction

\( c \): Index represents crops

\( c_1 \): A parameter through which the operation and maintenance costs of the dam are accounted for as a percentage of the construction cost

\( c_2 \): A factor associated with the O&M cost per unit area of the irrigation district

\( f_i(.) \): Probability density function of economic return of project \( i \)

\( s \): Index represents stochastic scenarios

\( x \): Attribute level of project return

\( y \): Index represents years

\( \gamma \): Discount factor

\( \lambda \): Parameter of EGCL order which indicates risk attitude of a decision-maker

\( \mu_i \): Mean of the economic return measure associated with project \( i \)

\( \sigma_i \): Standard deviation of the economic return measure associated with project \( i \)

**INTRODUCTION**

Irrigation planning and management under uncertainty, formulated within the framework of stochastic programming (SP), has been widely considered in the literature. Most often, existing literature considers uncertainty of the rainfall, reservoir inflow, and sometimes demand (evapotranspiration requirement) and is limited to systems of a reservoir and a single-crop irrigation district (de Lucia 1969; Dudley et al. 1971a, 1971b, 1972; Dudley & Burt 1973).

Stochastic dynamic programming (SDP) is among the models extensively used in irrigation planning and management under uncertainty. Such models have been used to deal with: (1) determination of optimum inter-seasonal water allocation and irrigation area in a reservoir-farm-aquifer system considering rainfall and streamflow...
uncertainty (de Lucia 1969); (2) determination of optimum intra-seasonal water allocation in a system with a known reservoir storage and irrigation area allowing for stochasticity of rainfall and irrigation requirement in a short-term framework (Dudley et al. 1971a); (3) determination of optimal seasonal acreage using a medium-term model (Dudley et al. 1971b); (4) determination of optimum irrigation area by a long-term model (Dudley et al. 1972); and finally, (5) integration of the problems related to optimal reservoir design and allocation policies with considerations regarding uncertainty of rainfall, runoff, and demand (Dudley & Burt 1973). SDP models, however, suffer from the curse of dimensionality problems, in which computational burden increases exponentially with the number of state variables (crops). Therefore, Dudley & Hearn (1993) considered the use of appropriate limited state variables representing a reservoir–irrigation district system in an SDP model, where the soil water content of the irrigation district was observed to be more significant than reservoir deficit (a state variable representing the deficit in reservoir storage to supply required soil water content of irrigation district) in a reservoir–irrigation district system.

When the number of crops increases, the strategy of selecting the state variables fails to work properly because for each crop, a state variable should be considered, which results in a large number of state variables that is not tractable by SDP models. Therefore, alternative approaches that integrate SDP with other models, e.g., dynamic programming (DP), linear programming (LP) and nonlinear programming (NLP), need to be used. In this regard, multi-crop models of stochastic irrigation planning and management have been used in determining optimum crop pattern and water allocation policies using SDP-LP (Dudley et al. 1976). Paudyal & Manguerra (1990) solved the same problem taking uncertainty of rainfall and streamflow into account through an SDP-DP model. Vedula & Mujumdar (1992) applied an SDP-DP model for optimizing reservoir operation and water allocation policies with a known crop pattern considering uncertainty of reservoir inflow, after that LP replaced DP to improve the computational efficiency (Vedula & Kumar 1996). Ghahraman & Sepaskhah (2002) used NLP instead of LP to account for nonlinearity of crop yield functions. SPs have also been solved using a Monte Carlo simulation (MC) approach in which a large number of realizations of stochastic processes and random variables are input to a deterministic optimization model (e.g., Cutore et al. 2008; Savic et al. 2009). To increase the computational efficiency, some studies have concentrated on using analytical methods such as first-order reliability method (e.g., Tolson et al. 2004; Babayan et al. 2005).

While the results of stochastic decision analysis depend on the type of the stochastic order considered (Tung & Yang 1994), all the above-mentioned models have used simple traditional stochastic orders of expected value and mean-variance. This is because traditional methods may fail in solving stochastic programs in the presence of stochastic orders other than simple expected value and mean-variance. Although there are a number of studies using non-traditional stochastic orders in agricultural and irrigation decision analysis (e.g., Anderson et al. 1977; Harris & Mapp 1986), they have not been presented in an SP framework. This paper presents different formulations of SP with respect to the type of stochastic order used in economic evaluation of a reservoir multi-crop irrigation project.

The solutions to the models presented in this paper will optimally determine the reservoir’s active capacity, area of irrigation district, and crop pattern as the most influential design parameters (variables) affecting benefit–cost evaluation of a reservoir–irrigation district project. The models are solved by a hybrid method integrating particle swarm optimization (PSO) and MC. The methodology includes the procedure of the probabilistic economic evaluation of the project that compares the project’s alternatives based on the adopted stochastic orders. Detailed introduction of the two stochastic orders known as stochastic dominance (SD) and expected gain-confidence limit (EGCL) for probabilistic economic evaluation of projects can be found in the following section. The section after that describes the problem of designing a reservoir multi-crop irrigation district under hydrological uncertainty, which is formulated as an SP model, and the solution approach, i.e., PSO–MC. The results of applying the proposed stochastic programs to the optimal design problem of the Hajiarab irrigation system, located in Ghazvin Province of Iran, are presented in the results and discussion section. The final section ends the paper with a summary and conclusions.
PROBABILISTIC PROJECT EVALUATION

The behavior of a decision-maker can be considered as risk aversion, risk neutral, or risk taking. Due to the fear of possible failure of an intended project performance caused by uncertainties, most public investment decision-makers, especially water-resource project managers, are risk averse (Tung & Yang 1994); thus, the procedure of project economic evaluation used in this study takes the risk averse behavior of decision-makers as an assumption.

There are a number of criteria and measures to evaluate the economic performance of a project, such as net present value (NPV), benefit-over-cost ratio (B/C), internal rate of return (IRR), etc., among which, NPV is used in this study. Consider maximization of NPV as the objective function of a project selection or ranking problem or an optimization problem in which NPV is a function of some stochastic variables. In this case, comparison of any two project alternatives or solutions requires the knowledge of the statistical properties of the NPV which is a function of decision variables. The frequently used statistical property of a project return that is characterized by NPV is the expected value of the project return by which different solutions are compared. However, the use of expected return may not account for the full extent of the uncertainty associated with the project’s return; hence more complete statistical information of an economic performance measure, carried by the variance, percentiles, or even the probability distribution function, than the economic performance measure, carried by the variance, percentiles, or even the probability distribution function, than the full extent of the uncertainty associated with the project return.

This feature in project evaluation might be an attribute of any economic performance measure, such as B/C or NPV. Project 1 dominates project 2 if the probability density function (PDF) of project 1, \( f_1(x) \), dominates that of project 2, \( f_2(x) \). This is called SD condition. In theory, the SD test can be performed to any degree involving multiple integrals of the difference in the cumulative distribution functions (CDFs) of the economic performance measure associated with two project alternatives.

In the first-degree stochastic dominance (FSD) test, no risk attitude of any type is assumed. The basic assumption is that a decision-maker prefers more of an attribute, \( x \) (such as NPV), to less. The FSD test states that project \( j \) dominates project \( i \) if

\[
\int_{-\infty}^{\infty} [f_i(x) - f_j(x)] dx = F_i(x) - F_j(x) 
\]

for all levels of attribute \( x \) with a strict inequality for at least some \( x \). Also, \( F_i \) represents the CDF of the economic measure for project \( i \). In the case of the FSD test being indecisive, the two projects are non-dominant or efficient. Then the relative preference of the two projects can further be tested by the second-degree stochastic dominance (SSD) test. In the SSD test, the procedure further assumes that the decision-maker is risk averse. Based on the SSD test, project \( j \) dominates project \( i \) if

\[
\int_{-\infty}^{\infty} [F_i(x) - F_j(x)] dx 
\]

for all levels of attribute \( x \) with a strict inequality that holds for some \( x \). The SSD test involves an evaluation of integration of CDF of the two project returns. Again, if the SSD test is indecisive, the third-degree stochastic dominance (TSD) test can be conducted which further assumes that the risk aversion of the decision-maker diminishes as the return gets larger and the attitude towards risk could not drastically change from risk aversion to risk taking. The decision rule of the TSD test is that project \( j \) dominates project \( i \) if and only if, \( E(X_j) \geq E(X_i) \), and

\[
\int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} [F_i(x') - F_j(x')] dx' \right) dx 
\]

for all values of attribute \( x \).
In order to compare projects under uncertainty conditions, one can use a five-step procedure proposed by Tung et al. (1995). Suppose there are two projects $i$ and $j$ to be compared.

1. Feasibility test: A project is considered feasible if the mean value of NPV is non-negative. The requirement of non-negative expected NPV ensures that the aggregate cost for the project can be – at least in the long run – recovered by the aggregate benefit. The decision rule at this step is to eliminate those projects with negative expected NPV from further consideration.

2. FSD test: If both of the projects pass the feasibility test, the FSD test is conducted according to Equation (1). When the FSD test is indecisive and the two projects remain efficient, the SSD test is then performed if the decision-maker is risk averse.

3. SSD test: To carry out the SSD test, one can apply Equation (2) for all possible values of NPV based on which a decision is made about whether the two projects under consideration are efficient, or one project dominates the other. When a definite conclusion cannot be made, the TSD test can be applied for further evaluation.

4. TSD test: With an assumption of decreasing risk-aversion behavior of the decision-maker, one may use Equation (3) over all possible values of NPV to provide information about the relative dominance between the two projects.

5. Final test: It is possible that an evaluation procedure remains indecisive after applying the first four screening steps. To choose the most economically efficient project, one can make the decision by using the following rule:

$$\text{if } \Pr(\text{NPV}_i \geq \text{NPV}_j) \begin{cases} >0.5, & \text{project } i \text{ dominates project } j. \\ <0.5, & \text{project } j \text{ dominates project } i. \end{cases}$$

(4)

**Expected gain-confidence limit (EGCL)**

Based on the EGCL rule, project $i$ dominates project $j$ if $\mu_i \geq \mu_j$ and $\mu_i + \lambda \sigma_i \geq \mu_j + \lambda \sigma_j$ for a suitably chosen value for $\lambda$ and without holding both equalities simultaneously, where $\mu_i$ and $\mu_j$ are means of the economic return measure, e.g., NPV, associated with, respectively, projects $i$ and $j$ (Baumol 1965). Also, $\sigma_i$ and $\sigma_j$ are standard deviations of the return measure for the projects.

The sign of $\lambda$ indicates the risk attitude of a decision-maker. For a risk-averse decision-maker, $\lambda$ is negative valued, while it is zero for a risk-neutral decision-maker, and for a risk-taking person, $\lambda$ is positive valued. The larger the positive value of $\lambda$, the more aggressive (optimistic) the decision-maker will be.

Based on EGCL, one can use the following procedure when comparing project alternatives under uncertainty (Tung & Yang 1994):

1. Feasibility test: Same as the SD procedure.
2. If the two projects remain efficient after the feasibility test, one can evaluate them further based on EGCL criteria in which project $i$ dominates project $j$ if $\mu_i > \mu_j$ and $\mu_i + \lambda \sigma_i \geq \mu_j + \lambda \sigma_j$ or $\mu_i = \mu_j$ and $\mu_i + \lambda \sigma_i > \mu_j + \lambda \sigma_j$.
3. The same as the final test described in the SD procedure.

**IRRIGATION PLANNING AND RESERVOIR DESIGN**

**General stochastic program (GSP)**

The case study system considered in this paper includes a dam reservoir and a downstream irrigation district. The aim is to find the optimum values of reservoir capacity, area of the irrigation district, and crop pattern. These design variables – affecting the procedure of economic evaluation of the project – are considered as decision variables. Moreover, inflows to the reservoir are assumed to be random variables of a stochastic process, causing the economic return from the project to be uncertain. The problem can be formulated in the form of a general stochastic program (GSP). The GSP formulation is represented below:

$$\begin{align*}
\text{Max}_{\alpha} & \quad \text{NPV} = PVB - PVC \\
\text{s.t.} & \quad PVB_s = \sum_y \left\{ \frac{1}{(1 + \gamma)^{y-1}} IA_{s,y} \sum_{e} CP_{e,c} Y_{c} P_{c} \right\} \forall s,
\end{align*}$$

(5)-(6)
\[ \text{PVC}_s = \sum_y \left\{ \frac{1}{(1+\gamma)^y} IA_{s,y} \sum_c CP_c \cdot CPC_c \right\} + (a_1K + b_1)(1 + c_1) + a_2TIA(1 + c_2) \quad \forall s, \]

\[ \text{LB}_c \leq CP_c \leq \text{UB}_c \quad \forall c, \]

\[ \sum_c CP_c = 1, \]

\[ \text{TIA} \leq \text{MaxLand}, \]

\[ \text{Max} \quad IA_{s,y} \quad \forall s, y, \]

\[ \text{st.} \]

\[ S_{s,y,t+1} = S_{s,y,t} + Q_{s,y,t} - R_{s,y,t} \quad \forall t, \]

\[ BF_t + IA_{s,y} \times \left( \sum_c CP_c \times CD_{c,t} \right) / IE \leq R_{s,y,t} \quad \forall t, \]

\[ S_{s,y,t} \leq K \quad \forall t, \]

\[ IA_{s,y} \leq \text{TIA} \] (15)

The objective function of the stochastic program is the maximization – in the stochastic order sense – of NPV which is achieved by the project construction over a long-term horizon as expressed in Equation (5), where \( \Omega \) is the stochastic order. The expression of ‘maximization in the stochastic order sense’ represented by \( \text{Max} \) emphasizes the fact that the program aims to find the optimum solution(s) that dominates other solutions based on stochastic orders. \( \text{PVB} \) and \( \text{PVC} \) are, respectively, the vectors of present values of benefit and cost. These vectors consist of the outcomes of random variables of the present value of benefit and present value of cost, respectively.

The present value of benefit – based on crop yields and prices – is determined through Equation (6), where \( \gamma \) represents the discount factor. The revenue due to cultivation of crop \( c \) in year \( y \) is estimated by product of the price of crop per unit mass, \( P_{ec} \), actual crop yield per unit cultivated area, \( Y_c \), and the area under cultivation of the crop, \( IA_{s,y} \times CP_c \). Note that \( IA_{s,y} \) is the total irrigated area in year \( y \) of scenario (outcome) \( s \), and \( CP_c \) is a fraction representing the share of crop \( c \) in total cultivation area. As explained later, total cultivation area \( (IA_{s,y}) \) is determined for each year of each time series (scenario) of reservoir inflow, separately. Finally, the present value of gross benefit, \( \text{PVB} \), is determined considering a cash flow diagram for the benefit achieved during the project’s lifetime.

The present value of cost can also be determined through Equation (7), which incorporates three constituents. The first term on the right-hand side (r.h.s.) of Equation (7) is present value of production costs over the project life, where \( CPC_c \) is production cost of crop \( c \) per unit area. The second term represents capital and maintenance costs of dam construction as a linear function of reservoir capacity, \( K \). Note that \( a_1 \) and \( b_1 \) are, respectively, associated with the variable cost and the fixed cost terms for dam construction and \( c_1 \) is a constant accounting for operation and maintenance (O&M) cost as a percentage of the construction cost. The last represents capital and maintenance cost of the irrigation district as a linear function of its area, i.e., \( \text{TIA} \). Note that \( a_2 \) is a factor associated with the construction cost per unit area of the irrigation district, while \( c_2 \) represents a factor associated with its O&M cost. Finally, the present value of total cost, \( \text{PVC} \), is determined considering a cash flow diagram for different cost components during the project’s lifetime.

Crop pattern \( (CP) \), which is a vector consisting of elements of fraction of area under cultivation of a crop to the total irrigated area, is constrained by upper and lower bounds (inequality Equation (8)). The bounds, \( \text{LB} \) and \( \text{UB} \), are estimated based on socio-economic studies. It is obvious that the sum of elements of the crop pattern vector must be equal to unity (Equation (9)). Also, size (area) of irrigation district is finite and must be smaller than the total land available as stated by inequality (Equation (10)), where \( \text{Maxland} \) is the maximum irrigable land for cultivation.

The above stochastic program is a multi-year long-term optimization scheme which includes yearly short-term optimization schemes. In this program, uncertainty of reservoir inflows can be dealt with by using the MC approach,
incorporating several scenarios (time series) of reservoir inflow into the program. Each scenario will result in an outcome of other dependent random variables and stochastic processes of interest involved in the model formulation. The objective function of short-term optimization schemes (Equation (11)), whose value is evaluated separately for each year of each scenario (time series), is maximization of the irrigated area. The short-term schemes are, respectively, constrained by reservoir balance equations, irrigation water requirement, baseline flow, bounds on reservoir storages, and an upper bound on the irrigated area according to Equations (11) to (15), where $R_{s,y,t}$ and $Q_{s,y,t}$ are, respectively, volumes of water released from and entered into the reservoir in month $t$ of year $y$ of scenario $s$, and $S_{s,y,t}$ is reservoir storage at the beginning of month $t$. $CD_{c,t}$ is net irrigation water requirement at farm per unit area of crop $c$ in month $t$, $IE$ is an irrigation efficiency factor as the fraction of water volume that effectively supplies the crop water requirement to water volume withdrawn from the source. $BF_{t}$ is the baseline flow required to support the ecosystem water requirement at downstream of reservoir in month $t$. Other variables are already defined.

Inequality (Equation (13)) ensures that the water released from the reservoir in each time step must be greater than or equal to total irrigation water requirement plus baseline flow, while inequalities (Equation (14)) and (Equation (15)) limit the reservoir's storage volume and irrigated area to active reservoir capacity, $K$, and area of irrigation district, TIA respectively.

The short-term optimization schemes associated with different years of a scenario are connected by the reservoir balance constraints (Equation (12)): stating that the reservoir storage at the beginning of each year must equal the reservoir storage at the end of the previous year.

The presented framework is able to model different time horizons that should be considered in irrigation planning. Although the design variables, i.e., the area of irrigation district, the reservoir capacity, and crop pattern, are estimated for a long horizon (usually a 50-year period), the cultivation area under irrigation could vary in each year depending on hydrologic conditions. The long-term optimization scheme determines the long-run decision variables, while the short-term optimization schemes determine the best values of state variables of irrigated area in each year, monthly releases from the reservoir, and the amounts of water allocated to different crops.

**Methodology**

The goal of the study is to evaluate the effects of streamflow uncertainty and the type of stochastic order on optimal design of a surface reservoir-irrigation district system using the GSP presented above. Accordingly, four mathematical programs are derived including:

1. Deterministic: This program, which can help us better understand the role of inflow uncertainty, includes only one scenario (50-year monthly time series) of reservoir inflow. As a result, the program will result in only one outcome of the dependent random variable (NPV). The program is called deterministic because it does not consider the uncertainty of streamflow process, so the uncertainty of the resulting net return from the project (NPV) is neglected.

2. Stochastic-EXP: Uncertainty of the reservoir inflow process causes the variables of the reservoir's storage and release, yearly irrigated area, benefit, production cost, and consequently NPV to become random variables. To take the uncertainty issue into account using the MC approach, the presented GSP is solved for a large number of inflow scenarios, each of which will generate a realization of NPV. In this program, the stochastic order of expected value (EXP) of NPV is used in evaluating the objective function values and comparing different solutions in the GSP.

3. Stochastic-SD: This program is different from stochastic-EXP only in the adopted stochastic order. To evaluate dominance relation between each of two alternatives (solutions), a stochastic order based on SD – as described earlier – is used in this case.

4. Stochastic-EGCL: The stochastic order used in this program is based on EGCL while other properties are the same as those of stochastic-EXP.

**Solution approaches**

All the stochastic programs explained above are nonlinear programs. In solving the above nonlinear programs, two
solution approaches are assessed. The first approach is PSO–MC, which is a combination of PSO algorithm and MC. To evaluate how well this approach performs, we compare it with the well-known gradient-based solution approach of nonlinear optimization, i.e., successive linear programming (SLP) (Bazaraa et al. 1993).

Particle swarm optimization (PSO)

The PSO method is a member of the evolutionary computation techniques that belong to the wide category of swarm intelligence methods for solving global optimization problems proposed by Kennedy & Eberhart (1995). It has been also used in several water-resource applications (e.g., Mousavi & Shourian 2000a, 2000b). In this algorithm, a population (swarm) of potential solutions to the problem under consideration is used to probe the search space using some individuals (particles) that all have an adaptable velocity (position change), according to which it moves in the search space.

Suppose that each particle is a candidate solution equivalent to a point in a D-dimensional space, so the i-th particle can be represented as $X_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,D})$. Therefore, each particle flies through the search space, depending on two important positions $P_i = (p_{i,1}, p_{i,2}, \ldots, p_{i,D})$, which is the best position that particle $i$ has found so far ($p$-best), and $P_g = (p_{g,1}, p_{g,2}, \ldots, p_{g,D})$, which is the global best position ($g$-best) which has been identified in the entire population. The change of $i$-th particle's position is given by its velocity $V_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,D})$. Equation (16) updates the velocity for each particle in the next iteration step, whereas Equation (17) updates each particle’s position in the search space:

$$
V_{i,d}^{t+1} = x(\omega V_{i,d}^t + c_1 r_1^t (p_{i,d}^t - x_{i,d}^t) + c_2 r_2^t (p_{g,d}^t - x_{i,d}^t)) \quad (16)
$$

$$
x_{i,d}^{t+1} = x_{i,d}^t + V_{i,d}^{t+1} \quad (17)
$$

where $d = 1, 2, \ldots, D$; $i = 1, 2, \ldots, N$ and $N$ is the size of the swarm; $x$ is called constriction factor which is used in constrained optimization problems in order to control the magnitude of the velocity; $\omega$ is called inertia weight; $c_1$, $c_2$ are two positive constants, called cognitive and social parameters, respectively; $r_1$, $r_2$ are random numbers uniformly distributed in [0,1]; and $n = 1, 2, \ldots, n_{\text{max}}$, represents the iteration number.

The PSO algorithm, which is one of the potential approaches of global optimization, was selected in this study to be combined with serially linked linear programs within a MC framework to solve the presented nonlinear stochastic programs. Regarding the parameter values of the PSO, experimental results advise that the initial value of the inertia weight, $\omega$, be large enough, in order to promote global exploration of the search space, and decrease it gradually to get more refined solutions (Shi & Eberhart 1998). Therefore, an initial value around 1.2 and a gradual decline towards 0 was used in this study. Parameters $c_1$ and $c_2$ were taken equal to 0.5 based upon the suggestions provided by Parsopoulos & Vrahatis (2002). The constriction factor, $x$, was also set to 0.9.

PSO–MC approach

The PSO–MC solution approach consists of two modules of PSO and MC. The serially linked yearly short-term optimization schemes are solved by revised simplex algorithm (Bixby 1992) within a MC framework, whereas the long-term scheme (the main part of GSP) is solved by the meta-heuristic PSO algorithm. The PSO–MC method takes advantage of meta-heuristic optimization techniques in handling nonlinearity of functions and, at the same time, the computation efficiency of LP solvers.

Before starting the PSO–MC approach, historical records of monthly reservoir inflow are taken, and a suitable probability distribution is fitted to data of each month. Using the fitted probability distributions, one hundred 50-year time series (scenarios) of monthly inflow are generated. The generated time series are input to the PSO–MC approach.

At each iteration of PSO algorithm, different candidate solutions are generated based on the location and velocity equations (Equations (16) and (17)). Meanwhile, the objective function value of each candidate solution can be evaluated only after solving serially linked yearly linear programs. This is because $I_{A_p}$ values in Equations (5) to (7) (values of total area under irrigation in each year) are the variables reflecting the effect of stochastic reservoir inflow and reservoir operation on the objective function. Also as stated before, the objective function of GSP, i.e., NPV, is a
random variable; therefore by means of each inflow scenario, a realization of objective function’s value – i.e., a realization of NPV random variable – is obtained. Upon repeating the same procedure within the MC module for all the generated scenarios (times series) of inflow process, a large number of realized values of NPV result for each candidate solution, from which the required statistics of NPV can be estimated. For stochastic-EXP, only the expected value of NPV is evaluated while for stochastic-EGCL, the standard deviation is also estimated. For stochastic-SD, the CDF of NPV is estimated. The statistics and CDF of NPV, depending on the stochastic order used in each of the programs, are used subsequently to check the optimality (objective function evaluation) of the candidate solutions. This procedure leads to the selection of the solutions (particles) of interest, $p$-best and $g$-best, at each iteration of the PSO module and continues until the PSO stopping criterion is met.

Figure 1 illustrates the flow diagram of the PSO–MC approach, schematically. In this approach the PSO is the master part within which MC is called for objective function evaluation of the candidate solutions. As described in the section PSO, PSO is an iterative algorithm by which the feasible space is searched to find the best or near-best solution. To do so, several candidate solutions are randomly generated first which are improved then in subsequent iterations of the algorithm based upon their objective function values. This means that the objective function value of each candidate solution must be evaluated. For objective function evaluation, the MC procedure, which is about solving serially linked yearly linear programs, is performed for determining the values of the variables used in the objective function. In other words, the MC procedure is embedded in the PSO for objective function evaluation of each candidate solution.

It is worth mentioning that the constraints of the GSP associated with the short-term optimization schemes are satisfied through the simplex method. All other constraints are linear (constraints (8) to (10)). The bound inequalities (8)
and (10) are met while generating the values of candidate solutions within the bounds in the first iteration of PSO. Then the feasibility of the solutions are checked in the next iterations of the PSO, where CP and TIA values are modified in case of any constraint violation. Therefore, a repairing strategy has been adopted for constraint handling in the PSO.

RESULTS AND DISCUSSION

Case study

Hajiarab agricultural area located in Ghazvin Province of Iran faces water scarcity in supplying agricultural water requirements. The largest part of water demand in this region is supplied by groundwater resources; hence the groundwater level has been gradually decreasing in recent years. To prevent aquifer degradation, groundwater exploitation should be considerably reduced. As a result, construction of the Balakhano Dam on the Hajiarab River has been planned in order to supply water to the downstream irrigation district. The plan’s objective is to maximize the economic benefit resulting from the agricultural production while sustaining surface and groundwater resources (Abanpajoh Consulting Eng. Co. 2004a).

Table 1 presents the crop pattern and the data used as input to the optimization model. The yield of each crop has been estimated based on FAO guidelines (FAO 1998), while price and production costs and other input data have been collected from agro-economic studies (Abanpajoh Consulting Eng. Co. 2004b). Irrigation efficiency was estimated to be about 36% (Abanpajoh Consulting Eng. Co. 2004c). Construction and maintenance costs for the dam and the irrigation district systems were estimated based on the system’s characteristics and the economic data available (Abanpajoh Consulting Eng. Co. 2004c). The values of constant coefficients in Equation (7) were obtained as: \( a_1 = 899 \), \( b_1 = 500000 \), \( c_1 = 0.6 \), \( a_2 = 250 \), and \( c_2 = 3 \), where the Iranian Rial (IR) is the monetary unit. Discount factor was considered equal to 6%.

A 30-year streamflow record was available at a hydro-metric station near the Balakhano Dam’s reservoir entrance point. After comparing the goodness-of-fit measures for three types of probability distributions of normal, log-normal, and exponential, a log-normal distribution was selected and fitted to historical data of each month. The distributions were then used to generate 100 scenarios of 50-year monthly inflow time series. This means that \( 100 \times 50 = 5000 \) linear programs are solved in the MC module for each candidate solution being evaluated by PSO. Table 2 compares the statistical characteristics of historical and generated data.

Deterministic program

Table 3 presents the values of different constituents of NPV for the optimum solution of the deterministic program. The gross benefit achieved from irrigation in the region is about two times larger than the associated net benefit. In other words, the cost of the project and the net benefit achieved are approximately the same.

The significance of the irrigation district area among the other decision variables is clear, as the largest component of the cost is related to the production cost, which mainly depends on the irrigation district area. Also, the benefit value depends on the revenue gained from agriculture which is directly related to the irrigation district area.

The deterministic program was solved by two different solution approaches of SLP and PSO–MC. As expected, the SLP approach was faster than the PSO–MC with an objective function value \( 1.2004 \times 10^8 \) very close (a difference less than 0.1%) to that of the PSO–MC \( 1.1996 \times 10^8 \). This

<table>
<thead>
<tr>
<th>CP bounds (%)</th>
<th>CP bounds (%)</th>
<th>Maximum yield (ton/ha)</th>
<th>Price (1,000 Rials/ha)</th>
<th>Total cost (1,000 Rials/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crop</td>
<td>Upper</td>
<td>Lower</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>35</td>
<td>25</td>
<td>11</td>
<td>750</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>11</td>
<td>15</td>
<td>750</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>5</td>
<td>4</td>
<td>1,200</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>13</td>
<td>30</td>
<td>500</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>7</td>
<td>2.5</td>
<td>3,500</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>4</td>
<td>15</td>
<td>750</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>5</td>
<td>30</td>
<td>550</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>8</td>
<td>25</td>
<td>600</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>2</td>
<td>4.5</td>
<td>1,500</td>
</tr>
</tbody>
</table>
shows the suitability of the PSO–MC in finding a good or a near-global solution to the problem.

Despite the closeness of the best objective function values of the two approaches, their decision and state variables’ values were different. The biggest difference resulting from the two approaches was in the variable of irrigation district, while their optimal capacities of the dam were almost the same. Table 4 presents more details in this regard. Results suggest constructing the Balakhanlo Dam with a capacity equal to 50 million cubic meters. Also, about 30% of the possible cultivated area is irrigated by the surface reservoir while the remaining part of the total cultivated area should be supplied by ground water resources.

Some interesting observations on optimum crop pattern can be seen. Optimum crop pattern depends on the interaction between the net benefit per hectare (difference of production benefit and cost) and the irrigation water requirement of each crop. In Table 4, crops with a smaller net benefit or water demand, i.e., Crop 3, Crop 7, Crop 8, Crop 9, and Crop 6, are cultivated as much as possible with $CP_c$ values close to their maximum bounds. Because of improper conditions of these two factors, Crop 1 and Crop 5 have been kept at their minimum (lower bound) cultivation area. Crop 2 is modest because of its mediocre situation in terms of the two mentioned factors.

### Number of hydrologic scenarios

A sensitivity analysis was carried out to evaluate the effect of the number of inflow scenarios (time series) on the statistics of the objective function, i.e., expected value and standard deviation of NPV. Figure 2(a) shows the results: the curves of expected value of NPV were sketched with respect to the decision variable of the district area. The curves associated with the different number of samples are slightly different. Although a 250-scenario case, whose result is almost the same as a 1,000-scenario case, is more preferable,
a 100-scenario case could also be acceptable considering the computational burden of the PSO–MC approach. In spite of the fact that the expected values of NPV for 250-scenario and 1,000-scenario cases are very close, the cases are more different with respect to standard deviation of NPV (Figure 2(b)). One can see a similar trend in the variation of standard deviation while the area of irrigation district increases.

### Comparison of solution approaches

As stated in the previous section, both SLP and PSO–MC approaches managed to solve the deterministic program and their optimum objective function values were nearly the same. To test and compare capabilities of these approaches to solving the stochastic programs – which are more complex than the deterministic one – the stochastic-EXP program was solved first by SLP. The program consisted of 125,417 variables and 205,127 constraints, compared with the deterministic program with 1,270 variables and 1,881 constraints. The SLP solver was not able to find any solution and after spending many iterations failed to solve the problem. Seifi & Hipel (2001) showed how incorporating multiple inflow scenarios in a long-term reservoir operation planning with stochastic inflows leads to a very large deterministic equivalent model which is hard to solve using conventional optimization methods. They presented an efficient interior-point optimization algorithm for solving the resulting equivalent problem by taking advantage of the specific structure of the problem.
The PSO–MC approach was, therefore, used to solve the stochastic programs. Table 5 presents the results. Column 2, pasted from Table 4, refers to the deterministic program for comparison purposes. Results of the stochastic-EXP program are given in column 3. Given that all 100 hydrologic scenarios were embedded in the stochastic programs, the problem size and the computational load of the PSO–MC algorithm were much more than those of the deterministic problem.

Table 5 | Comparison of the results of different programs

<table>
<thead>
<tr>
<th>Models</th>
<th>Deterministic</th>
<th>Stochastic-EXP</th>
<th>Stochastic-SD</th>
<th>Stochastic-EGCL (\lambda = 1)</th>
<th>Stochastic-EGCL (\lambda = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dam active capacity (mcm)</td>
<td>50.00</td>
<td>50.00</td>
<td>49.81</td>
<td>49.43</td>
<td>49.63</td>
</tr>
<tr>
<td>Crop pattern (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crop 1</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Crop 2</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>Crop 3</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Crop 4</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Crop 5</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Crop 6</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Crop 7</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Crop 8</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Crop 9</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Irrigation district size (ha)</td>
<td>2,383</td>
<td>1,834</td>
<td>1,685</td>
<td>1,594</td>
<td>1,905</td>
</tr>
<tr>
<td>Expected value of NPV (IR)</td>
<td>–</td>
<td>(1.172 \times 10^8)</td>
<td>(1.169 \times 10^8)</td>
<td>(1.172 \times 10^8)</td>
<td>(1.173 \times 10^8)</td>
</tr>
</tbody>
</table>

The PSO–MC approach was, therefore, used to solve the stochastic programs. Table 5 presents the results. Column 2, pasted from Table 4, refers to the deterministic program for comparison purposes. Results of the stochastic-EXP program are given in column 3. Given that all 100 hydrologic scenarios were embedded in the stochastic programs, the problem size and the computational load of the PSO–MC algorithm were much more than those of the deterministic problem.

**Comparison of the types of stochastic order**

Table 5 represents the results of different programs. As presented in the table, the results of the stochastic-EXP program are close to those of the deterministic program; but stochastic-EXP has resulted in a smaller irrigation area than that resulting from the deterministic program. Values in the third and the fourth columns of Table 5 clearly show the extent of the difference lying in the use of SD order instead of the simple expected value. Columns 3 and 4 show the difference is limited to a slight reduction in irrigation district area (about 8%), while no other significant change occurs in other variables. Meanwhile, the expected values of NPV of the programs are different. The solution of the stochastic-SD program with a smaller expected value of NPV dominates the solutions of stochastic-EXP and deterministic programs.

Figure 3(a) compares the CDFs of the NPV resulting from the best solution of stochastic-EXP and stochastic-SD. Although the CDFs are very close, the solution of

Figure 3 | (a) Comparison of CDFs of the objective function of the stochastic-EXP and stochastic-SD programs for their best solution; (b) graphical representation of SSD decision rule.
stochastic-SD dominates that of stochastic-EXP. A careful evaluation of the stochastic orders used for determining the two solutions shown in Figure 3(a) reveals that stochastic-SD’s solution dominates that of stochastic-EXP based on the SSD test. To better understand the SSD decision rule, two symbolic alternatives, i.e., $i$ and $j$, with their respective CDFs of NPV, i.e., $F_i(.)$ and $F_j(.)$, are shown in Figure 3(b). As discussed in Equation (2), the SSD test involves an evaluation of integration of the CDFs of the two projects’ (alternatives) returns (NPVs). The r.h.s. of the equation represents the cumulative difference between the areas under the two CDFs. The area under $F_i(.)$ bounded by $F_j(.)$ is denoted by $A$, while $B$ is the area under $F_j(.)$ bounded by $F_i(.)$. Consequently, when $A$ is larger than or equal to $B$, project $i$ is dominated by project $j$ in the SSD sense. Although the CDFs of NPV associated with stochastic-SD’s best solution and that of stochastic-EXP are very close, their relative status is similar to that of alternatives $j$ and $i$ illustrated in Figure 5(b). It is worth mentioning that the difference between the two mentioned areas, i.e., $A-B$, is equal to 0.00113434 ($10^8$ thousand Rials). Therefore, the stochastic-SD’s best solution dominates that of stochastic-EXP based on the SSD decision rule. For further details on graphical interpretation of SSD, refer to Tung et al. (1993).

Similar differences were found when comparing the results of the program that uses the EGCL order with those of other programs. For example, the EGCL program has resulted in a 30% reduction in the irrigation district area compared with the stochastic-EXP program. The stochastic-SD and stochastic-EGCL programs behave more similarly with a 5% difference in their irrigation areas.

The sensitivity analysis was performed to assess the effect of parameter $\lambda$ on the EGCL results as presented in columns 5 and 6 of Table 5. For $\lambda = 10$, the decision-maker is more of a risk-taker compared with when $\lambda = 10$; therefore, the size of irrigation district has increased by about 20%. In other words, a more risk-taking decision-maker will choose a larger area of the irrigation district.

As mentioned before, the irrigation district area was the most sensitive variable among the others. The expected value and the standard deviation of the objective function (NPV) is depicted in Figure 4(a). As can be observed, the expected value sharply increases up to 1,200 ha, then flattens and stays rather the same for values in the range of 1,200–1,700 ha and finally decreases very gently. The behavior of standard deviation cannot be distinctively defined. It forms a zero-value region for 0–1,000 ha followed by a fast increasing part and a slow increasing region afterwards. It finally terminates with a plateau associated with a constant value equal to $7.8 \times 10^6$ IR. This explains why the more risk-taking behavior produces greater optimal values for the area of irrigation district. By increasing the area of irrigation district the variance of NPV increases; thus, a more risk-taking decision-maker who thinks positively about the variance of NPV, between the projects with the same expected value of the NPV, will select the project associated with the bigger variance of NPV. On the other hand, for the values of the district area greater than 1,200 ha, an increase of the district area decreases the expected value of NPV. Therefore, the optimal value of the district area is determined based on the trade-off between expected value and standard deviation of NPV.

![Figure 4](https://iwaponline.com/jh/article-pdf/15/2/591/387033/591.pdf)  
**Figure 4**: Variation of expected value and standard deviation of objective function (a) and sensitivity analysis with respect to parameter $\lambda$ in EGCL program (b).
Figure 4(b) illustrates the direct relation between the two variables defined as variation of $\mu + \lambda \times \sigma$, an indicator parameter for objective function in the stochastic-EGCL program, and the size of irrigation district.

It is worth mentioning that execution run time of the PSO–MC solution approach for solving stochastic-SD took about 20 h using a normal PC (with a Quad CPU, Q8400 @ 2.67 GHz speed and 3.5 GB of RAM).

SUMMARY AND CONCLUSIONS

This study explored the effects of three aspects: (1) uncertainty of streamflow, (2) the use of stochastic orders other than a simple expected value, and (3) the solution approach, i.e., a gradient-based technique compared with a hybrid solution approach, on the optimal design of a reservoir–multi-crop irrigation district system. Four nonlinear mathematical programs of deterministic, stochastic-EXP, stochastic-SD, and stochastic-EGCL were developed in this regard. The second program uses the simple expected value order, whereas the third and fourth employ the stochastic orders of the SD and the EGCL, respectively.

The programs were used for optimal design of Hajiarab Irrigation District located in Ghazvin Province of Iran. Design variables were the area of irrigation district, the dam reservoir capacity, and the crop pattern. Two approaches of SLP and PSO–MC, a combination of PSO and serially linked linear programs within a MC framework were used to solve the programs. SLP was able to solve only the first program of deterministic type, while PSO–MC could solve all the programs, although reaching the global optimum may not be assured.

The case study results showed that the gross benefit achieved was approximately two times greater than the net benefit. It also revealed the agricultural production cost as the most important term among the cost components. The results suggested construction of a dam with maximum possible storage capacity. However, the reservoir could supply water for only 1,594–2,383 ha, i.e., 16–23% of the total potential land available in Hajiarab (10,000 ha); therefore, the rest had to be irrigated using other resources. The most sensitive variable to the type of stochastic order was the area of irrigation district. This was due to its significant effect on the expected present value of the net benefit used as the objective function.

The stochastic programs deploying hydrologic uncertainty could reduce the size (area) of the irrigation district up to 23–30% when risk-aversion attitude was considered. Moreover, the use of stochastic orders based on SD concepts, rather than a simple expected value order, reduced the size of the irrigation district further (8%). This reduction was in excess of the amount that resulted from replacing the deterministic program with a stochastic program with the stochastic order type of expected value. The increase of parameter $\lambda$ in the stochastic-EGCL program, by which the risk-aversion attitude of the decision-maker is controlled, from one to ten caused the size of the irrigation district to improve by 20% as the decision-maker becomes more risk-taking.

It is worth mentioning that the maximization of an economic objective function was considered in this study. However, the importance of considering other criteria (social, environmental, etc.), which can be approached by multi-objective optimization, should be highlighted for future contributions.

ACKNOWLEDGEMENT

The help provided by Fariborz Jolai of Tehran University College of Engineering and J. Amini is greatly acknowledged.

REFERENCES


Downloaded from https://iwaponline.com/jh/article-pdf/15/2/591/387033/591.pdf by guest


First received 19 July 2012; accepted in revised form 23 October 2012. Available online 6 December 2012.