Ann Oper Res (2014) 223:309–328
DOI 10.1007/s10479-013-1455-8

Stochastic vs deterministic programming in water management: the value of flexibility

Yousaf Shad Muhammad · Georg C. Pflug

Published online: 14 October 2014

Abstract In the paper we develop a two stage scenario-based stochastic programming model for water management in the Indus Basin Irrigation System (IBIS). We present a comparison between the deterministic and scenario-based stochastic programming model. Our model takes stochastic inputs on hydrologic data i.e. inflow and rainfall. We divide the basin into three rainfall zones which overlap on 44 canal commands. Data on crop characteristics are taken on canal command levels. We then use ten-daily and monthly time intervals to analyze the policies. This system has two major reservoirs and a complex network of rivers, canal head works, canals, sub canals and distributaries. All the decisions on hydrologic aspects are governed by irrigation and agricultural development policies. Storage levels are maintained within the minimum and maximum bounds for every time interval according to a power generation policy. The objective function is to maximize the expected revenue from crops production. We discuss the flexibility of two stochastic optimization models with varying time horizon.

Keywords Indus Basin Irrigation System · Stochastic modeling in water systems · Water resources management · Stochastic vs deterministic programming

1 Introduction

The Indus Basin Irrigation System (IBIS) is the largest contiguous irrigation system in the world and has been developed over the last 140 years. The Indus river basin stretches from

This research work is sponsored by Higher Education Commission (HEC), Pakistan.

Y.S. Muhammad (✉)
Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan
e-mail: yousaf@inbox.com

Y.S. Muhammad
e-mail: yousaf@rcms.nust.edu.pk

G.C. Pflug
University of Vienna, Universität Strasse 5/3, 1010, Wien, Austria
e-mail: georg.pflug@univie.ac.at

the *Himalayan Mountains* in the north to the dry alluvial plains of Sind, in the south, with an area of 944,574 km^2 (see World Commission on Dams 2000). The vast irrigation system in Pakistan is comprised of two major storage reservoirs, 19 barrages (canal head work), 44 main canals (or command areas) with a conveyance length of 57,000 kilometers, and 89,000 water courses with a running length of more than 1.65 million kilometers. This vast irrigation system covers more than 18 million hectares of irrigated land in Pakistan, a country with the highest irrigated to rain-fed land ratio in the world (4:1).

Pakistan depends on irrigation and water resources for 90 percent of its food and crop production (see World Bank 1992).

Agriculture is the back bone of its economy.

This system has been developed at a significant financial cost. Therefore, it necessitates sophisticated scientific water management policies. Many advanced countries have developed the methods based on system sciences, operation research and mathematical modeling for decision making, in order to provide successful water resource management.

Decision-making of reservoir release for irrigation, when it is being operated under a power generation policy, involves subtle considerations regarding the nature of the crops being irrigated and timing related to decision for these. Determining the amount of water released from a reservoir as complex as the Indus networks must be supported by a comprehensive mathematical decision-making mechanism. It is necessary to consider the crop water requirement along with the competition with other crops, especially when water resources are scarce.

1.1 Stochastic programming in water systems

The process of determining the best allocation and utilization of available scarce resources, is as old as man himself. The uncertainty of future water resources adds additional complexity to the problem of optimum allocation. This allocation problem has been studied by economists, engineers and mathematicians for centuries. However, over the last four decades, it is being studied in the context of stochastic optimization.

In deterministic LP, all components are considered to be known. In practice, this rarely happens. Consider the transition of reservoir storage from one volume in one period to a different volume in the next. The transition results partially from the release of water for various uses, which can be controlled, and partially from inflow into the reservoir and reservoir losses such as evaporation and seepage, which can not be controlled. Therefore, the first component can be made deterministic, but the last two components cannot because they are random acts of nature. Inclusion of such random components defines the LP formulation as stochastic linear programming (SLP).

The stochastic dynamic programming (SDP) has proved to be a potential tool in developing reservoir operation models in the past (see Butcher 1971; Torabi and Mobasheri 1973; Dudley and Burt 1973; Mawar and Thorn 1974; Roefs and Guitron 1975; Bogle and O'Sullivan 1979; Oven-Thompson et al. 1982; Stedinger et al. 1984; Karamouz and Houck 1982; Kelman et al. 1982). Yeh (1985) presents a comprehensive state-of-the-art review of the various reservoir operation models. Most of the approaches in reservoir management for irrigation treated seasonal crop water demand for irrigation as deterministic, variability in reservoir inflow was taken into account (see Hall et al. 1969; Schweig and Cole 1968). Some exceptions to this generalized approach (see Sanford 1969; Dudley 1970 and Burt and Stauber 1971) incorporate stochastic water demand but assume a deterministic water supply. More citations toward both stochastic water demand and stochastic inflow are considered in Dudley et al. (1971, 1976), Dudley (1970, 1972) and Dudley and Burt (1973).

Due to uncertainty in the random behavior of hydrologic variables, single decision- making mechanisms for reservoir operation and crop water allocation are addressed by using SDP (see Houghtalen and Loftis 1988; Dudley 1988; Dudley and Scott 1993; Vedula and Mujumdar 1992; Vedula and Kumar 1996; Ravikumar and Venugopal 1998). LP, Deterministic Programming (DP) and SDP are used in seasonal and intra-seasonal allocation of deficit water, competing crops and crop yield optimization (see Paudyal and Gupta 1990; Rao et al. 1990; Azar et al. 1992; Mannocchi and Mecarelli 1994; Sunantara and Ramirez 1997; Paul et al. 2000; Anwar and Clarke 2001). LP, DP and simulation are effective tools in adaptive operations, such as real time forecasts of hydrologic variables (see Dariane and Hughes 1991; Rao et al. 1992; Mujumdar and Ramesh 1997; Wardlaw and Barnes 1999).

Many studies address uncertainty due to randomness of hydrologic variables and single decision-making mechanisms for reservoir operation and crop water allocation by using stochastic dynamic programming (see Houghtalen and Loftis 1988; Dudley 1988; Dudley and Scott 1993; Vedula and Mujumdar 1992; Vedula and Kumar 1996; Ravikumar and Venugopal 1998). A two-phase Stochastic Dynamic Programming (SDP) model was developed by Umamhesh and Sreenivasulu (1997) for optimal operation of irrigation reservoirs under a multi-crop environment. In the first phase they maximized the release from the reservoir and in the second phase they minimized the deficit of water when different crops were competing for scarce water resources. A (state of the art) review of SDP is presented in Labadie (2004) and references herein specially for reservoir operations. A stochastic formulation is presented in Ganji et al. (2006) for crop water demand when there is a deficit in irrigation water.

1.2 A two-stage stochastic program for IBIS

A classical two-stage SLP model with fixed recourse (see Dantzig 1955; Beale 1955) is defined as

$$\min \quad c^\top x + E_\xi\big[q(\xi)^\top y(\xi)\big], \tag{1}$$

$$\text{s.t.} \quad \mathbf{A}x = b, \tag{2}$$

$$\mathbf{T}(\xi)x + \mathbf{W}y(\xi) = h(\xi), \tag{3}$$

$$x \geq 0, \quad y(\xi) \geq 0, \tag{4}$$

where c and b are known vectors, $\mathbf{A}$ is a known matrix and $\mathbf{W}$ is assumed as a fixed recourse matrix. ξ denotes randomness, which represents the possible scenarios. $\mathbf{T}$ is a stochastic matrix and h is random right hand side. E_ξ represents the mathematical expectation with respect to ξ.

In our IBIS scenario-based stochastic model, the first-stage decision variable x is the vector of crop area to be sown in every canal command. The random variable ξ represents the rainfall and water inflow. The realization of ξ is unknown at the time when x is to be determined, but becomes known when the recourse decision y is to be made. The recourse decision variable $y(\xi)$ is the vector of crop area cultivated within each canal command when actual scenario ξ becomes known. Notice that the first-stage decisions x are, however, chosen by taking their future effects into account on the recourse decisions in the second stage.

The above formulation for the IBIS scenario-based stochastic model can be presented as to maximize the expected revenue from crop sales after subtracting the initial fixed cost and the cost of labor. Other expenditures are ignored such as water cost (a public property, farmers get their share for free in proportional to their land holdings), farm rent (farmers are the owners of their land) and initial cost (includes seed cost, fertilizer, etc.).

In comparison to the general model above, our model is characterized by the fact that constraint (3) of the recourse model is very simple: The right hand side is zero and the matrices **T** and **W** are identity matrices.

To be more precise, the second stage constraint is

$$x - y(\xi) - z(\xi) = 0, \tag{5}$$

$$x, y(\xi), z(\xi) \geq 0, \tag{6}$$

where $z(\xi)$ is a slack variable, represents the area dropped from cultivation. In different notations, the constraint is

$$0 \leq y(\xi) \leq x, \tag{7}$$

i.e. we can cultivate only what we have sown.

This paper is organized as follows. Section 1 explains the model construction for the Indus Basin Irrigation System (IBIS). In Sect. 2, we discuss how the scenarios for the model were generated. Section 3 explains the power generation process under a working policy. The results and discussions are given in Sect. 4, and summary in Sect. 5.

2 Model description

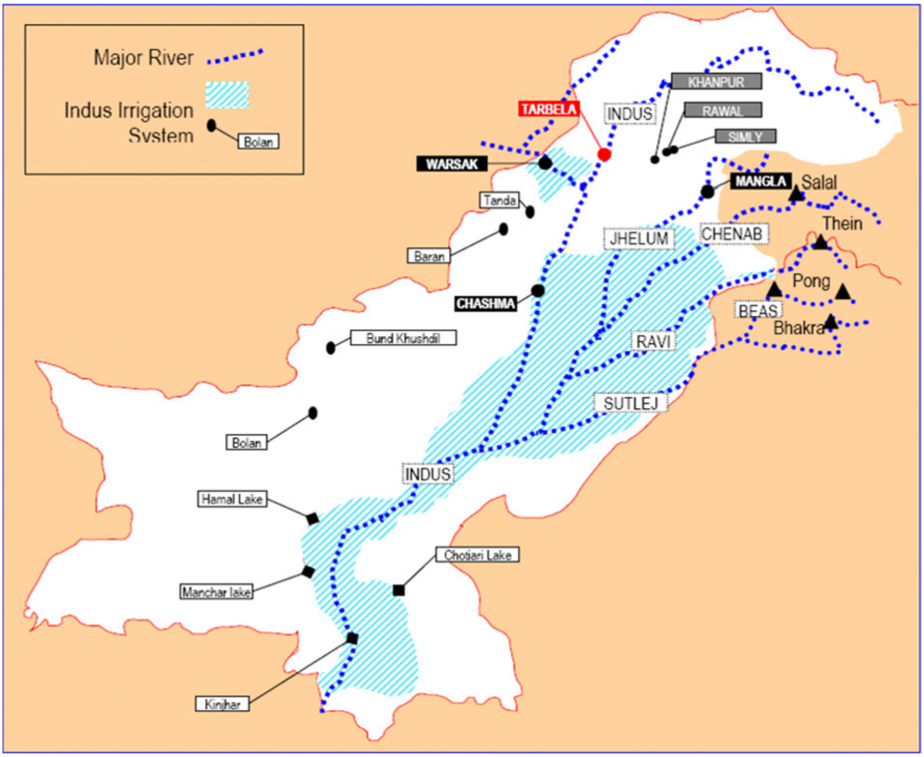
IBMR (Indus Basin Model Revised) is a mathematical model developed by Water And Power Development Authority (WAPDA) (2007) in collaboration with the World Bank. It is a mathematical model for irrigation water management implemented in the Indus Basin Irrigation System (IBIS), which deals with all decisions regarding reservoir operations, down stream cropping, import/export policies. The inputs into the model are deterministic and incorporate a yearly decision-making mechanism which refreshes its decision on a monthly basis. In reality, the inputs including stream flow (inflow), rainfall, evaporation, crop response and yield, etc. are random processes. One month time interval is too large when rapid changes in agro-climatic conditions and inflow into the system are expected.

Earlier models developed for optimal reservoir operation for irrigation deal with different aspects of this problem with varying degrees of complexity. The Stochastic Optimization approach explicitly incorporates uncertainties due to the randomness of hydrologic variables. Developing a stochastic model for this system, which samples stochastic inputs over shorter time intervals, may enhance the system's efficiency.

We present a ten-daily and a monthly time interval model, based on scenarios for this basin. The present model addresses policy related issues and considers the following features.

- **Complex network involved:** As discussed earlier, this system involves a very big rivers (7 in all), several nodes (34 in all), node-to-node links and multiple demand sites (44 in all). Figure 1 shows the Indus River System in the study area.

  A hydrologic agro-economic model was studied by Ximing et al. (2003). The model extends integration of a management water supply system and irrigation farming system to a spatially large and complex system. This, however is not a stochastic model. The model was developed for a river basin network, including multiple nodes and demand locations (6 demand sites).



Source: WCD (Terbela Dam a Scoping Report)

Fig. 1 Indus Irrigation System (Source: WCD, Terbela Dam a Scoping Report)

- **Stochastic inflow and rainfall:** As discussed earlier, irrigation water management involves several stochastic parameters. In this study, we restrict ourselves to stochastic inputs for inflow and rainfall only. For the sake of computational difficulties, the remaining parameters are adjusted to their average values.
- **Two-stage stochastic model:** We developed a two-stage stochastic programming model for IBIS. The first-stage decision variables (here and now) include the vectors (crops area) that are sown in different canal commands. These decisions must consider the uncertainty in future realization of the scenarios. The second stage decision variables (wait and see) are the vectors (crops area) that are cultivated within different canal commands when the actual scenarios become known.

The model is developed in order to determine the optimal cropping patterns and irrigation scheduling for the described system. In the formulation of the model, the following factors are considered.

– It is a 1-year planning model.
– Operating policy for hydrologic scheduling and cropping occurs on ten-daily (i.e. every month is divided into 3 ten-daily intervals) and monthly intervals.

– Canal l belong to exactly one zone, say $z(l)$. Zone-wise canals are as follow

Zone	Canals (l)
Zone1	2a, 03, 04, 05
Zone2	11, 12, 13, 14, 18, 22, 23, 24, 25, 26, 27
Zone3	01, 2b, 06, 07, 08, 09, 10, 15, 16, 17, 19, 20, 21, 28, 29, 30 to 43

– Land occupation (cropping calender) is available for each crop and each canal command (e.g. for canal command 5 in Fig. 3).
– The year is further divided into two seasons, the wet season (April to September) and the dry season (October to March). Scenarios are generated within each season from their distribution of total inflow as well as for each rainfall zone (see next section).
– Crop water requirements within each canal command are computed by using the results of a study in the basin (see KaleemUllah et al. 2001).
– Other information on crop inputs such as labor, fertilizer, seed cost, crop yield etc. is extracted from documents of the WAPDA of Pakistan, Statistical Bureau of Pakistan (2005) and Economic Survey of Pakistan (2005).
– Data on ground water, system inflow, evaporation from storage and canals, data about system networks, and area for each canal command, are available from WAPDA Pakistan and IRSA Pakistan.
– Reservoirs are operated under a power generation policy (i.e., maintaining the storages volume within minimum and maximum bounds over time steps), see Eq. (16) and Fig. 6. The water scheduling policy for the reservoirs aims at optimizing the agriculture production. Power generation is not an objective, but the technical complications for the power generation enter into the optimization model as constraints.
– The indexes and index sets are in Table 1.

Table 1 Index and Cardinality

Index	Name	Index set	Cardinality
c	crop	C	13
i	storage	N_S	2
k	state	K	4
l	canal command	L	44
n	node	N	35
r	river	R	7
s	scenario	S	200
t	time-step	T	month = 12, ten-daily = 36
z	rainfall zone	Z	3

Note: the node set consists of three types of nodes i.e. storage nodes or reservoirs (N_S), rim station nodes (N_R) and ordinary nodes (N_O). The node set is $N = N_S \cup N_O \cup N_R$

– **Deterministic data**

$\Delta_{i,min}$	minimum water volume at storage i (km^3)
$\Delta_{i,max}$	maximum water volume at storage i (km^3)
H	risk free upper bound capacity (volume) of system infrastructure (km^3)
p_c	price of crop c (dollars/10^3 kg)
$Share_k$	percentage share of state k in aggregate surface water
$Labor_z$	labor force available in zone z (hours)
$CropArea_{zc}$	maximum area suitable for crop c in zone z (10^6 hectares)
$Land_l$	land resources of canal l (10^6 hectors)
ε_{it}	evaporation from storage i during time t (km^3)
λ_{cl}	average production of crop c in canal l (10^3 kg/hectare)
ϑ_{cl}	initial (sowing) cost for a unit land on crop c in canal l(dollars/hectare)
δ_l	carrying to field efficiency of canal l
ν_{clt}	water needed for crop c in zone z during time t (millimeter)
β_{clt}	labor hours for crop c incanal l during time t (hours/hectare)
θ_{cl}	total labor cost for a unit land of crop c in canal l (dollars/hectare)
a_{clt}	indicator for crop c in zone z during time t (present $=1$, absent $=0$)
τ_{lt}	ground water available in canal l during time t (km^3)
$LinkCap(n, n'')$	capacity of link $(n, n'') \in \mathbb{N}$ (km^3)
$CanalCap_l$	capacity of canal l (km^3)

– **Stochastic data and scenarios probability**

$\tilde{\sigma}^s_{zt}$	rainfall in zone z duringtime t under scenario s (millimeter)
$\tilde{\alpha}^s_{rt}$	inflow in river r during time t under scenario s (km^3)
π^s	probability of scenario s
ω^s_{int}	outflow from storage i to node n during time t under scenario s (km^3)
$\eta^s_{n'nt}$	release of water towards node n from node n' during time t under scenario s (km^3)

– **Decision variables**

X_{lc}	area sown for crop c in zone z canal l (hectares)
Y^s_{lc}	area cultivated for crop c inzone z canal l underscenario s (hectares)
W^s_{lt}	water released in canal l zone z duringtime t underscenario s (km^3)
Δ^s_{it}	water level at storage i during time t under scenario s (km^3)

– Objective function is to maximize the revenue from crops in the basin.

2.1 Objective function

The objective function seeks to maximize the expected revenue from all crops. Every crop $c \in \mathrm{C}$ is possible in all canal commands $l \in \mathrm{L}$ but has different yields, say λ_{cl}. The sale price of crop c is p_c with θ_{cl} representing the total labor cost of cultivation for a unit (hectare) of land for crop c in canal command l. First-stage decision variable X_{lc} is the area in hectares, sown for crop $c \in \mathrm{C}$ in canal command $l \in \mathrm{L}$. The second stage decision variable is a scenario-based variable Y^s_{lc}, which is the area cultivated under scenario $s \in \mathrm{S}$. The objective is to maximize the expected total return

$$\max\left[\sum_{s\in S}\pi^s\left\{\sum_{l\in L}\sum_{c\in C}(p_c\lambda_{cl}-\theta_{cl})Y^s_{lc}\right\}-\sum_{l\in L}\sum_{c\in C}\vartheta_{cl}X_{lc}\right]. \tag{8}$$

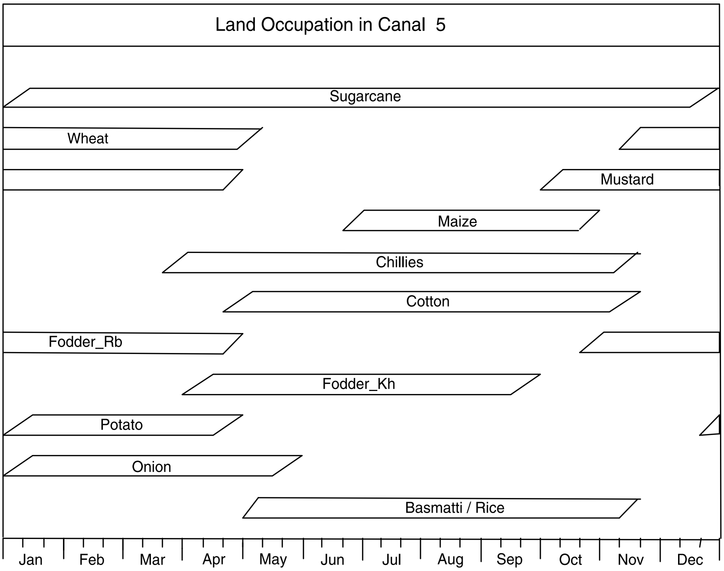


Fig. 2 Land occupation a_{clt}

2.2 Constraints area balance

The total utilization of area within every canal should never be greater than the area available over time horizon t. Different crops have different time schedules for occupying the field, and sowing, cultivation and harvesting follow one after the other. The parameter a_{clt} is an indicator which signifies the occupation of a field for crop c in canal l over time horizon $t \in T$. $a_{clt} = 1$ when a crop is present in the fields (at any stage: sowing, cultivation or harvesting) and 0 otherwise. The sum of the area being utilized (sown or cultivated) for all crops should not be more than the area $Land_l$ in a particular canal command.

$$\sum_{c \in C} a_{clt} X_{lc} \le Land_l; \quad \forall\, l \in L,\ t \in T \tag{9}$$

The following variable upper bound ensures that the area cultivated is less then the area sown in all possible scenarios.

$$X_{lc} - Y_{lc}^s \ge 0 \quad \forall\, l \in L,\ c \in C,\ s \in S \tag{10}$$

There is an upper bound on cropwise area used within each zone. Every crop is not sown more than the area designated within all zones. Farmers grow some crops within their land holding such as fodder for cattle rather than buying it from the local market. Moreover, it is difficult to transport such crops from one zone to another due to transportation costs and loss in quality during transport. Due to soil types and characteristics, the whole area in a canal command, is not suitable for every crop, but only a particular size area can be used for a certain crop. If $CropArea_{zc}$ is the upper limit on the area to be sown in zone z crop c (see Paudyal and Gupta 1990), then

$$\sum_{l:z(l)=z} X_{lc} \le CropArea_{zc} \quad \forall\, z \in Z,\ c \in C \tag{11}$$

Data about $Land_l$ and $CropArea_{zc}$ were taken from WAPDA, Pakistan.

2.3 System network

The Indus river system involves a complex system network. There are several nodes, node-to-node links and canals. The set of nodes is $N = N_S \cup N_O \cup N_R$ and the set of links in the system is $\mathbb{N} \subseteq N \times N$.

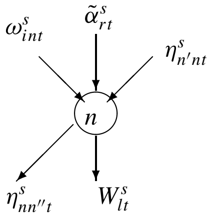
Let $\eta^s_{n'nt}$ represent the trafficking of water towards node n from node n' (where $n' \in N_O$) during time t under scenario s. The total water trafficking towards node n, from other nodes n' is $\sum_{n':(n',n)\in\mathbb{N}} \eta^s_{n'nt}$. And the divergence of water to other nodes n'' (where $n'' \in N_O$) from any node n is $\sum_{n'':(n,n'')\in\mathbb{N}} \eta^s_{nn''t}$.

Rim station nodes (i.e. nodes where some river joins the system) also receive direct inflow from uninterruptible rivers. Every node $n \in N_R$ receive inflow from river $r(n)$. Direct uninterruptible inflow at node $n \in N_R$ during time t under scenario s is denoted by $\tilde{\alpha}^s_{r(n)t}$.

The other source of trafficking-in of water at any node is outflow from a storage. Let ω^s_{int} denote the outflow from storage i (where $i \in N_S$) to node n during time t under scenario s. Then, the total trafficking of outflow towards any node n can be expressed as $\sum_{i:(i,n)\in\mathbb{N}} \omega^s_{int}$.

The important function at canal head work (node) of this system, is to regulate canal diversions, in an optimal way. Each canal l diverts water exactly from one node $n(l)$. If W^s_{lt} is the amount of water release into canal l during time t under scenario s, then the total water diverted into canals from node n can be expressed as $\sum_{l:n(l)=n} W^s_{lt}$.

The water balance at node n can be shown as



Combining all the trafficking of water at a node, we have

$$\sum_{n':(n',n)\in\mathbb{N}} \eta^s_{n'nt} + \sum_{i:(i,n)\in\mathbb{N}} \omega^s_{int} + \tilde{\alpha}^s_{r(n)t} - \sum_{n'':(n,n'')\in\mathbb{N}} \eta^s_{nn''t} - \sum_{l:n(l)=n} W^s_{lt} \geq 0; \quad \forall\, n \in N,\ t \in T,\ s \in S \tag{12}$$

Water in the links and in the canals should not exceed by their capacities. $LinkCap(n, n'')$ is the capacity of link $(n, n'') \in \mathbb{N}$ and $CanalCap_l$ is the capacity of canal l. Then

$$0 \leq \eta^s_{nn''t} \leq LinkCap\big(n, n''\big); \quad \forall\, t \in T,\ s \in S,\ \big(n, n''\big) \in \mathbb{N} \tag{13}$$

$$0 \leq W^s_{lt} \leq CanalCap_l; \quad \forall\, l \in L,\ t \in T,\ s \in S \tag{14}$$

2.4 Water linkage

Let Δ^s_{it} denote the water volume stored in reservoir i at the end of time t under scenario s. Let ε_{it} denote the *fixed* evaporation from storage i during time t. Storages are rim stations as well. Each storage i receives some inflow from exactly one river, say $r(i)$. The water linkage constraint can be expressed

$$\Delta^s_{i(t-1)} + \tilde{\alpha}^s_{r(i)t} - \Delta^s_{it} - \sum_{n:(i,n)\in\mathbb{N}} \omega^s_{int} - \varepsilon_{it} = 0; \quad \forall\, i \in N_S,\ t \in T,\ s \in S \tag{15}$$

where ω^s_{int} is the outflow from storage i towards node n at time t under scenario s. $\tilde{\alpha}^s_{r(i)t}$ is the inflow of river r at time t under scenario s that falls into the prescribed storage i.

Storage volume is maintained according to the system infrastructure (and storage/power generation policy). If $\Delta_{it,min}$ denotes the minimum volume at time t in storage i and $\Delta_{it,max}$ denotes the maximum volume at time t in storage i, then

$$\Delta_{it,min} \le \Delta^s_{it} \le \Delta_{it,max}; \quad \forall\, i \in N_S,\ t \in T,\ s \in S \tag{16}$$

2.5 Water balance

We follow a deterministic crop demand policy from Schweig and Cole (1968) and Hall et al. (1969). The amount of water required for a unit of land of crop c in canal l during the time t is denoted by υ_{clt} and $\tilde{\sigma}^s_{zt}$ denotes the rainfall in the zone z during the same time t under scenario s. The difference $(\upsilon_{clt} - \tilde{\sigma}^s_{zt})$ is provided by the irrigation source. If W^s_{lt} is the amount of water put into the irrigation canal l with carrying-to-field efficiency δ_l, and τ_{lt} is the ground water available in canal command l during time t, then the water consumption balance can be expressed as

$$\delta_l W^s_{lt} + \tau_{lt} - \sum_{c\in C} \max\bigl(\upsilon_{clt} - \tilde{\sigma}^s_{z(l)t}, 0\bigr) Y^s_{lt} \ge 0; \quad \forall\, l \in L,\ t \in T,\ s \in S \tag{17}$$

Let ω^s_{int} denotes the outflow from the storage node $i \in N_S \subseteq N$ to node n during the time t under scenario s, and $\tilde{\alpha}^s_{r(n)t}$ is the inflow (or sum of all rivers inflow that fall in node n) in river $r(n) \in \mathrm{R}$ during time t under scenario s. The water consumption should be bounded below the water available in the system, i.e. the aggregate water diversion in the canals should not be greater than the water available from direct inflow and outflow from storages.

$$\sum_{r:n\in N_R} \tilde{\alpha}^s_{r(n)t} + \sum_{i:(i,n)\in\mathbb{N}} \sum_{n:(i,n)\in\mathbb{N}} \omega^s_{int} - \sum_{l\in L} W^s_{lt} \ge 0; \quad \forall\, t \in T,\ s \in S \tag{18}$$

2.6 Flood aversion constraint

Seasonal and annual river inflows in the Indus river system are highly variable (see Warsi 1991; Kijine and Van der Velde 1992; Ahmad 1993). Sixty percent of inflow may occur in only July and August. This river system has the potential to generate floods. We incorporate a flood aversion constraint in the present model to avoid loss of lives and property in down stream areas due to flooding. We kept the total surplus amount of water in the system below a risk free upper bound (H km^3) according to the system infrastructure. Moreover optimal storage operations are necessary as we do not have control over uninterruptible inflow. So, the constraint can be expressed

$$\sum_{r:n\in N_R} \tilde{\alpha}^s_{r(n)t} + \sum_{i:(i,n)\in\mathbb{N}} \sum_{n:(i,n)\in\mathbb{N}} \omega^s_{int} - \sum_{l\in L} W^s_{lt} \le H; \quad \forall\, t \in T,\ s \in S \tag{19}$$

2.7 Political constraint

The system serves several states which have a political accord among them, stating "*No state can use more than a stipulated percentage of the aggregate of total surface water available in a year*". If every canal l belongs to exactly one state $k(l)$ and $Share_k$ is the percentage share of state k, then

$$\sum_{l:k(l)=k}\sum_{t\in T} W_{lt}^{s} = Share_k \sum_{l\in L}\sum_{t\in T} W_{lt}^{s}; \quad \forall\, k,\ s \in S \tag{20}$$

2.8 Labor constraint

Labor constraint ensures the availability of labor within a zone for an optimal cropping pattern. If $Labor_z$ is the labor available in zone z and β_{clt} is the labor required for crop c in canal l over time horizon t, then

$$\sum_{l:z(l)=z}\sum_{c\in C} \beta_{clt} Y_{lc}^{s} \le Labor_z; \quad \forall\, z \in Z,\ t \in T,\ s \in S \tag{21}$$

2.9 Non negativity of the decision variables

$$X_{lc}, Y_{lc}^{s}, \Delta_{it}^{s},\ W_{lt}^{s} \ge 0 \tag{22}$$

3 Scenario generation in IBIS

In the present model, the inflows as well as the rainfall are treated as stochastic parameters, so that quantities which are unknown at the first decision stage, can be observed prior to the second decision stage.

In particular, we considered five random parameters: inflow for season 1, inflow for season 2, rainfall in zone 1, rainfall in zone 2 and rainfall in zone 3. Historical data have been used to calibrate the scenario model. Inflow data for season 1 and season 2 for the period of 1922–1923 to 2003–2004 were used to calibrate a normal distribution for the inflow data for each season. Figure 3 shows historic inflow and Gaussian fit. In a similar manner, a normal model for the rainfall per zone was estimated from historical data. Inflow in season 1 and season 2, shows an insignificant correlation with rainfall in the three zones and a marginal significant correlation between the two seasons. For simplicity, we ignored this correlation.

In contrast, monthly evaporation rates are considered deterministic and set to their average values.

Since the complexity of the optimization model increases with the number of scenarios, we used a backward scenario reduction technique to bring the number of scenarios down to 200.

To represent the stochastic inflows, we consider a discrete distribution sitting on just five points. The scenario generation technique is adopted from Pflug (2001) and Hochreiter and Pflug (2007), which aims at minimizing the approximation error.

If G is some continuous distribution function on $\mathbb{R}$, we replace this by a discrete distribution $\tilde{G}$ in such a way that the Wasserstein-distance d_1 between G and $\tilde{G}$ is minimal. Here the Wasserstein distance is defined as

$$d_1(G,\tilde{G}) = \sup\left\{\int f(u)dG(u) - \int f(u)d\tilde{G}(u) : L_1(f) \le 1\right\} \tag{23}$$

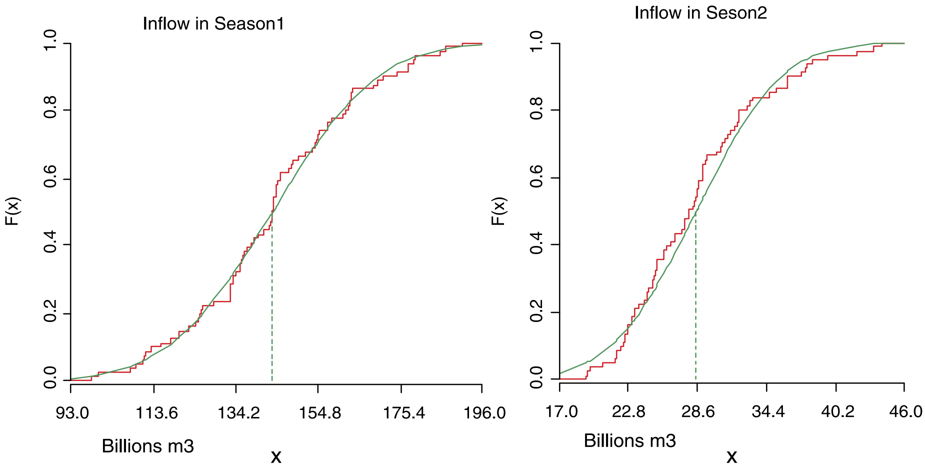
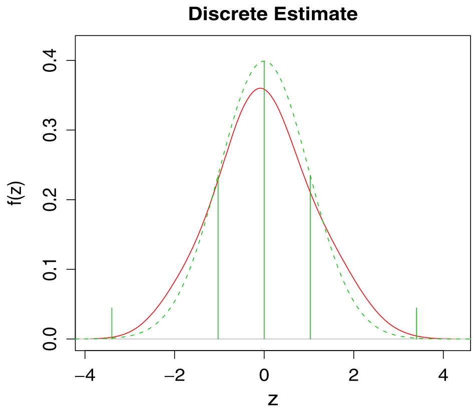


Fig. 3 Aggregative seasonal inflow (km^3) along with Gaussian model comparison

Fig. 4 Discrete estimate for inflow distribution along with historic inflow



where $L_1(f)$ is the Lipschitz-constant of f.

It is well known that this distance is related to the mass transportation problem (see Monge 1781; Rachev 1991). If G is the standard normal $N(0, 1)$ distribution, the optimal location of approximating mass points are -3.4, -1.029, 0, 1.029, 3.4 with masses 0.0446, 0.2589, 0.3930, 0.2589, 0.0446 respectively. Figure 4 shows the location of five points and their masses in the standard normal distribution. If $\tilde{G}$ sits on two points, the optimal z values are -0.7979, 0.7979 with masses that are 0.5, 0.5 respectively. If the distribution is $N(\mu, \sigma^2)$, these points have to be transformed by $\mu + \sigma z$ (see Pflug 2001). The total set of scenarios is generated by taking all possible independent combinations using the product probabilities (Table 2).

One might argue that 200 scenarios may not reflect the total variability of the stochastic parameters. However, the mere introduction of a scenario model, makes the solutions much more robust, compared to the deterministic case (see the comparison below). Moreover, the scenarios were chosen in such a manner that they represent the probability distribution quite well despite the small number of scenarios.

Table 2 Set of Scenarios

Variable	Number of scenarios
inflow season 1	5
inflow season 2	5
rainfall zone 1	2
rainfall zone 2	2
rainfall zone 3	2
total number of scenarios	$5 \times 5 \times 2 \times 2 \times 2 = 200$

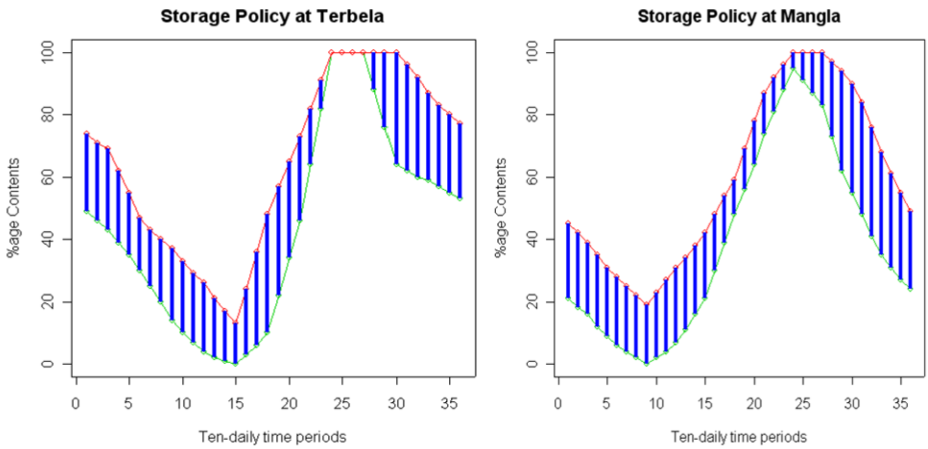


Fig. 5 Storage level upper and lower limits according to power generation policy i.e. the system will store water into these two storages (Terbela and Mangla) to generate a minimum required power from the system

4 Power generation policy in IBIS

Power generation is not an optimization objective here, but is a side benefit of water management. It is, however, interesting to analyze how optimal water management policies will influence the power generation (see Fig. 7).

Reservoirs in the network are operated according to a power generation policy i.e. stored water volume is maintained within the minimum and maximum bounds on each reservoir (see Eq. (16)):

$$\Delta_{it,min} \leq \Delta_{it}^{s} \leq \Delta_{it,max}; \quad \forall\, i \in N_S,\ t \in T,\ s \in S$$

In IBMR, target storage levels are set for each month. In the present model, we are also taking decisions on ten-daily time horizons. Accordingly, we linearly interpolate the target storage levels for all ten-daily periods. This storage policy is shown in Fig. 5.

4.1 Modelling conditions

We solve the model under three different modelling conditions (see Table 3).

1. A monthly time horizon with a single stage deterministic model, setting inflow and rainfall to their average values.

Table 3 Model dimensions

Model	Rows	Columns	Non zeros	Memory used	Time (sec)
Monthly deterministic	2773	2437	20405	<1 MB	4
Monthly stochastic	572,374	487,861	3,080,807	253 MB	2,280
Ten-daily stochastic	767,830	1,191,061	7,751,040	611 MB	20,429

Table 4 Cropping policy and revenue (Area 1000 hectares)

Area Sown->	Deterministic	Stochastic (monthly)	Stochastic (ten-daily)	Actual (2003–2004)
Crop	$\sum_l X_{lc}$	$\sum_l X_{lc}$	$\sum_l X_{lc}$	–
Basmatti	577	697	519	–
Irri	758	758	838	2503
Maize	81	81	87	896
Mustard	310	310	331	244
Sugarcane	924	924	1026	947
Fodder-kharif	1313	1313	1403	Not available
Fodder-rabi	1352	1352	1445	Not available
Cotton	2881	2881	3028	3221
Gram	880	880	468	1038
Wheat	2296	3624	4403	8330
Potato	32	32	77	111
Onion	56	56	60	122
Chilies	78	69	85	39
Total	11538	12977	13770	–
Exp. Rev. (bn USD)	5.318	6.130	7.175	Not available

2. A monthly time horizon two-stage stochastic model with 200 scenarios.
3. A ten-daily time horizon two-stage stochastic model with 200 scenarios.

The elapsed time, number of constraints and number of variables are given below. We solve all these models with the dual-simplex algorithm. We use CPLEX 9.1.2 as solver. CPU resources used to get all these solutions are Intel T-2300 1.66 GHz Core Duo Processor with 2 GB RAM.

5 Results and discussion

5.1 Cropping pattern

Model uses only 14.87 million hectares (mha) of land resources. This is the area under canal commands. In "Actual" column of the Table 4, reported area includes cultivated area from all resources i.e. 14.87 mha canal irrigated, 3.91 mha of other irrigated resources and about 13.06 mha of rain fed area (Economic Survey, Pakistan 2004–2005). In comparison, the actual area and results of the model indicate the difference in rice, wheat, maize and gram areas.

Rice represents the total of two varieties (Basmatti and Irri). Maize is a major crop in rainfed areas along with wheat and gram. Premature maize is used as fodder for cattle. It is also considered as fodder crop throughout the year, and farmers cultivate it two to three times each year. All the rainfed areas are used for wheat cultivation which are not the part of our land resources. That is why, the wheat area is almost twice in 'actual 2003–2004' comparing with all solutions. In rainfed areas, average wheat yield, is two to three times lesser than the irrigated areas. Gram is the most suitable crop for rainfed areas due to its low water requirement across the cultivation calender.

Some of the land might have been used twice or even more times. It is according to the optimal strategy the model might have adopted, strictly according to cropping calender (Fig. 2) as you see land can be used twice, wheat-maiz, wheat-fooder and wheat-cotton etc. More over, solutions given above threw light how they change land resources utilization from deterministic to stochastic models and than shows increase in flexibility by increasing area in ten-daily horizon model comparing with monthly horizon model as both models are two-stage models. This is because of flexibility created by shorter time horizon land occupation strategy. We tried this computation strategy by increasing more scenarios, but results were not much improved. This might be due to insufficient crepitation in the region and insufficient but volatile inflow in the system. *Deterministic vs Stochastic Solutions*.

5.2 The value of stochastic solution (VSS)

The scenario-based Stochastic Linear Programming model produces better results for both the monthly and ten-daily time horizons and shows a significant improvement, in land utilization. This improvement is further enhanced with the ten-daily time horizon, as shown in the Table 4. This increase is because of flexibility of field allocation to different crops with shorter time horizon In terms of revenue earned, the results indicate:

- EV (expected value or mean value) solution is obtained by replacing the random variables by their expected values. EEV is then obtained, which is defined as the expected result of using the EV solution. The parameter EEV, measures how the mean value solution perform, allowing second stage decisions to be chosen optimally as a function of using the EV solution and randomness (see Birge and Louveaux 1997). EEV is 5.473 billion USD.
- The expected revenue (agriculture + power) with the monthly two-stage stochastic model corresponding to the recourse problem (RP) is 6.130 billion dollars.
- The value of stochastic solution is: VSS = RP − EEV = 6.130 − 5.473 = 0.657 bn USD

  We discussed the results for only one stochastic model i.e. monthly time horizon model. Results may be more intrusting for ten-daily time horizon model.

This comparison gives a measure of value when using the decision mechanism from deterministic programming to stochastic programming with monthly time horizon models. The value of stochastic solution is quite significant standing at 657 million dollars. It represents the cost of ignoring uncertainty when comparing all benefits.

5.3 Expected value of perfect information (EVPI)

The expected value of perfect information (EVPI), is the difference between the expected solution with Perfect Information about future scenario say (PI) and the here-and-now solution. We calculate the EVPI for the monthly time horizon model only. This exercise is also possible and may more interesting with the ten-daily time horizon model. We substitute

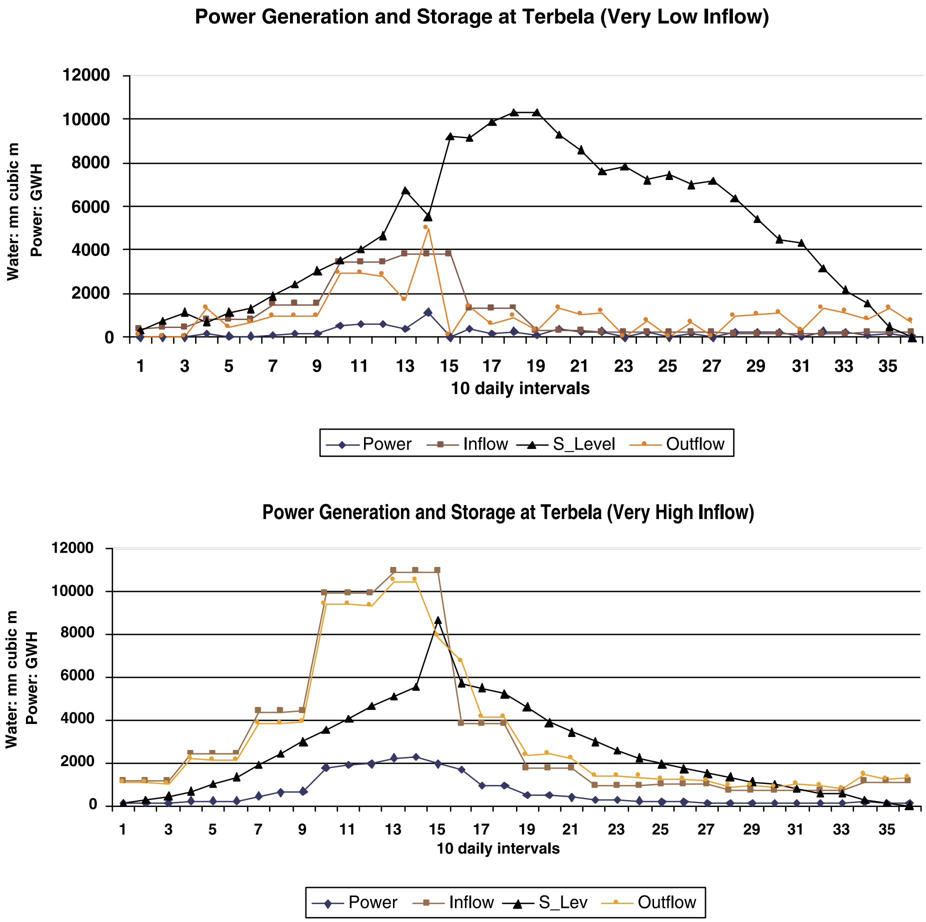


Fig. 6 Power generation and storage at Terbela in extreme scenarios

$X_{lc} = Y_{lc}^s$ in the model to get the Perfect Information (PI) solution. The perfect information (PI) solution, which considers benefits and losses, is the expected value of optimal solutions. It is 6.241 billion US dollars. The here-and-now solution corresponding to the recourse problem (RP) is 6.130 billion US dollars. The expected value of perfect information is given as:

$$\text{EVPI} = \text{PI} - \text{RP} = 6.241 - 6.130 = 0.111 \text{ bn USD}$$

This shows a difference of 111 million USD. This is the amount we are ready to pay to obtain perfect information about random parameters.

5.4 Storage and power generation

Pakistan consumed 57,491 GWH (giga watt hours) electric power in 2003–2004. The average production for the last five years (1999–2000 to 2003–2004) from hydro power generation was 21,085 GWH, which was 37 % of the total consumption (Source: Economic Survey, Pakistan 2004–2005). The power demand is projected to grow at an annual average rate of

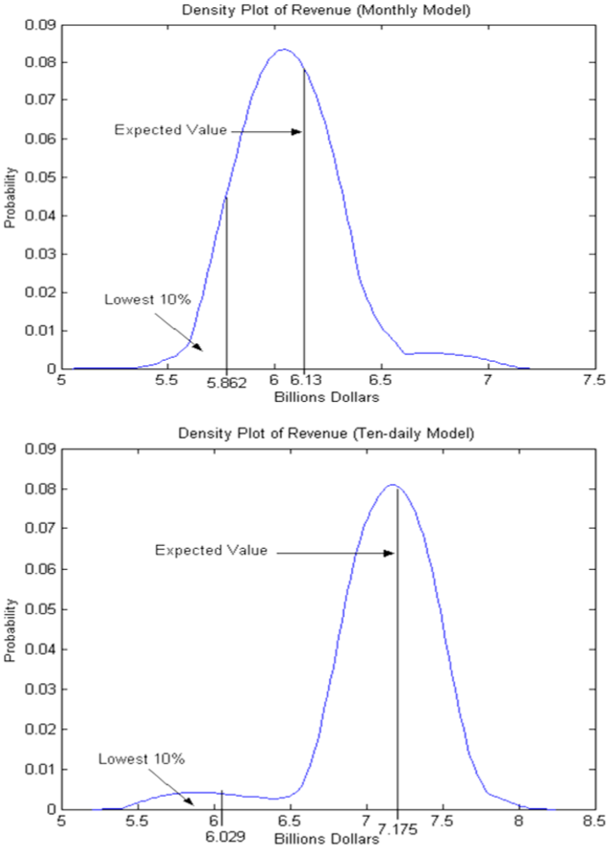


Fig. 7 Revenue plot (expected value and lowest 10 % are shown)

7.9 percent during next five years. With the available hydropower production capability, the ratio of hydro power production will continue to decrease unless the new water storages are established. With a ten-daily scenario based model, hydro power generation varies from 10,316 GWH to 29,352 GWH from a very low to very high inflow scenario. These results are consistent with the present power generation policy in the system. We run this model primarily for optimal cropping policy. Storage levels were not maintained to maximize power generation from the reservoirs, even then the model results regarding power generation performed well. Figure 7 shows inflow, outflow, storage level and power generated in extreme scenarios.

5.5 Revenue generation: the value of flexibility

Agriculture serves in Pakistan as the backbone of country's economy. 68 percent of the population directly or indirectly depend upon agriculture and (26 %) of the population lives below poverty line (Economic Survey, Pakistan 2004–2005). Most of this group are farmers living in rural areas where agriculture is their sole source of income. Any crisis in agriculture production, will hit this section of the population directly, and they will be more vulnerable in an adverse scenario. This alarming situation demands an appropriate implementation of hydrologic decision policies which will support this section of the population.

We solved the model under different modelling conditions. We focused on revenue generation by agriculture production. Here we observed that the ten-daily time horizon stochastic model gives a maximum expected revenue of 7.175 billion dollars compare to 6.130 and 5.318 with monthly stochastic and monthly deterministic models, respectively. If we look at the 10th decile of the revenue distribution over these scenarios, the ten-daily time horizon stochastic model has a higher 10th decile compared to the others. The revenue plot of the stochastic models is shown in Fig. 7. The 10th decile is the lower tail of the revenue distribution, representing extremely unlikely scenarios. This difference between results is due to the flexibility created by the ten-daily stochastic model in the land utilization over shorter time intervals (time horizon). The decision mechanism for shorter time intervals enhances the model's capability on managerial grounds. Moreover, managing hydrologic decisions for reservoir operations, subsequently, increase system's efficiency.

6 Summary

A stochastic programming model for the Indus Basin Irrigation System has been presented in this paper. It has been calibrated considering the complete network of rivers and canal system. This results in a huge model, specially, when it is operated under a ten-daily time horizon. In the stochastic model, we consider randomness in hydrologic variables, inflow and rainfall in the basin. The whole basin is divided into three rainfall zones. It can be used as an administrative tool for decision-making, formulating the cropping policies and scheduling the reservoirs release. We presented a comparison between stochastic and deterministic solutions. We also showed the flexibility a stochastic approach can have over a deterministic one. This flexibility increases, when we manage decisions over shorter periods of time. This gives a two fold comparison of different approaches. The value of a stochastic solution (VSS) and the expected value of perfect information (EVPI) illustrate the advantages of a stochastic approach over a deterministic one. Two hundred scenarios were generated for the hydrologic parameters. Although this is a small number of scenarios, the stochastic model leads to a significant improvement over deterministic one (see VSS). We incorporate a constraint avoiding flood over the maximum amount of surplus water. It restricts the surplus water below a certain level during all time periods according to system infrastructure. We provide a reconciliation among the provinces of Pakistan over surface water usage by incorporating a political constraint.

Acknowledgements We are greatly thankful to Dalia Bach from the University of Columbia, who helped us a lot to improve this paper. Thanks to referee for his value suggestions.

References

Ahmad, N. (1993). *Water resources of Pakistan*. Lahore: Miraj uddin Press.

Anwar, A. A., & Clarke, D. (2001). Irrigation scheduling using mixed-integer linear programming. *Journal of Irrigation and Drainage Engineering*, *127*(2), 63–69.

Azar, A. H., Murty, V. V. N., & Phien, H. N. (1992). Modeling irrigation scheduling for lowland rice with stochastic rainfall. *Journal of Irrigation and Drainage Engineering*, *118*(1), 36–55.

Beale, M. L. (1955). On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society. Series B. Methodological*, *17*, 17–184.

Birge, J., & Louveaux, F. (1997). *Introduction to stochastic programming*. New York: Springer.

Bogle, M. G. V., & O'Sullivan, M. J. (1979). Stochastic optimization of a water supply system. *Water Resources Research*, *15*(4), 778–786.

Burt, O. R., & Stauber, M. S. (1971). Economic analysis of irrigation in subhumid climate. *American Journal of Agricultural Economics*, *53*(1), 33–46.

Butcher, W. S. (1971). Stochastic dynamic programming for optimum reservoir operation. *Water Resources Bulletin (Harrisburg)*, *7*(1), 115–123.

Dantzig, G. B. (1955). Linear programming under uncertainty. *Management Science*, *1*, 197–206.

Dariane, A. B., & Hughes, T. C. (1991). Application of crop yield functions in reservoir operation. *Water Resources Bulletin (Harrisburg)*, *27*(4), 649–656.

Dudley, N. J. (1970). A simulation and dynamic programming approach to irrigation decision making in a variable environment. Agricultural Economics and Business Management Bulletin 9, Faculty of Agriculture Economics, University of New England, Armidale, Australia.

Dudley, N. J. (1972). Irrigation planning for optimal intraseasonal water allocation. *Water Resources Research*, *8*(3), 586–594.

Dudley, N. J. (1988). A single decision-maker approach to irrigation reservoir and farm management decision making. *Water Resources Research*, *24*(5), 633–640.

Dudley, N. J., & Burt, O. R. (1973). Stochastic reservoir management and system design for irrigation. *Water Resources Research*, *9*(3), 507–522.

Dudley, N. J., & Scott, B. W. (1993). Integrating irrigation water demand, supply and delivery management in a stochastic environment. *Water Resources Research*, *29*(9), 3093–3101.

Dudley, N. J., Howell, D. T., & Musgrave, W. F. (1971). Optimal intraseasonal irrigation water allocation. *Water Resources Research*, *7*(4), 770–778.

Dudley, N. J., Reklis, D., & Burt, O. R. (1976). Reliability, trade-offs and water resource development modeling with multiple crops. *Water Resources Research*, *12*(6), 1101–1108.

Economic survey of Pakistan (2005). *Govt. of Pakistan, Ministry of Finance*.

Ganji, A., Ponnambalam, K., Khalili, D., & Karamouz, M. (2006). A new stochastic optimization model for deficit irrigation. *Journal of Irrigation Science*, *25*, 63–73.

Hall, W. A., Tauxe, G. W., & Yeh, W. W. G. (1969). An alternate procedure for the optimization of operation for planning with multiple river, multiple purpose system. *Water Resources Research*, *5*(6), 1367–1372.

Hochreiter, R., & Pflug, G. (2007). Financial scenario generation for stochastic multi-stage decision process as a facility location problem. *Annals of Operations Research*, *152*, 257–272.

Houghtalen, R. J., & Loftis, J. C. (1988). Irrigation water delivery system operation via aggregate state dynamic programming. *Water Resources Bulletin (Harrisburg)*, *24*(2), 427–434.

KaleemUllah, M., Zaigham, H., & Saim, M. (2001). *Spatial distribution of reference and potential evaporation across the Indus Basin Irrigation System*. International Water Management Institute, Lahore. Working Paper 24.

Karamouz, M., & Houck, M. H. (1982). Annual and monthly reservoir operating rules generated by deterministic optimization. *Water Resources Research*, *18*(5), 1337–1344.

Kelman, J., Stedinger, J. R., Cooper, L. A., Hsu, E., & Yuan, S. Q. (1982). Sampling stochastic dynamic programming applied to reservoir operation. *Water Resources Research*, *26*(3), 447–454.

Kijine, J. W., & Van der Velde, E. J., Jr. (1992). *Irrigation management implications of Indus Basin climate change—case study*. Lahore: IIMI. (From "World commission on Dams" documents.)

Labadie, J. W. (2004). Optimal operation of multireservoir systems: state-of-the-art review. *Journal of Water Resources Planning and Management*, *130*, 93–111.

Mannocchi, F., & Mecarelli, P. (1994). Optimization analysis of deficit irrigation systems. *Journal of Irrigation and Drainage*, *120*(3), 484–503.

Mawar, P. A., & Thorn, D. (1974). Improved dynamic programming procedure and their application to water resource systems. *Water Resources Research*, *10*(2), 183–190.

Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Histoire de L'Academie des Sciences. Paris*, *29*, 352–362.

Mujumdar, P. P., & Ramesh, T. S. V. (1997). Real-time reservoir operation for irrigation. *Water Resources Research*, *33*(5), 1157–1164.

Oven-Thompson, K. I., Alercon, L., & Marks, D. H. (1982). Agricultural versus hydropower trade-offs in the operation of the High Aswan Dam. *Water Resources Research*, *18*(6), 1605–1614.

Paudyal, G. N., & Gupta, A. D. D. (1990). Irrigation planning by multilevel optimization. *Journal of Irrigation and Drainage Engineering*, *116*(2), 273–291.

Paul, S., Panda, S. N., & Kumar, D. N. (2000). Optimal irrigation allocation: a multilevel approach. *Journal of Irrigation and Drainage Engineering*, *126*(3), 149–156.

Pflug, G. (2001). Scenario tree generation for multi-period financial optimization by optimal discretization. *Mathematical Programming Series B*, *89*, 251–271.

Rachev, S. T. (1991). *Probability metrics and the stability of stochastic models*. New York: Wiley.

Rao, N. H., Sarma, P. B. S., & Chander, S. (1990). Optimal multicrop allocation of seasonal and intraseasonal irrigation water. *Water Resources Research*, *26*(4), 551–559.

Rao, N. H., Sarma, P. B. S., & Chander, S. (1992). Real-time adaptive irrigation scheduling under a limited water supply. *Agricultural Water Management*, *20*, 267–279.

Ravikumar, V., & Venugopal, K. (1998). Optimal operation of South Indian irrigation systems. *Journal of Water Resources Planning and Management*, *124*(5), 264–271.

Roefs, T. G., & Guitron, A. (1975). Stochastic reservoir model: relative computational effort. *Water Resources Research*, *11*(6), 801–804.

Sanford, S. C. (1969). Economics of water allocation within irrigation scheme. In *The economics of irrigation development: a symposium. Study 6*. Dep. of Agri. Eco. and Develop. Univ. of Reading, England

Schweig, Z., & Cole, J. A. (1968). Optimal control of linked reservoir. *Water Resources Research*, *4*(3), 479–497.

Statistical Year Book (Series). (2005). *Statistical Bureau of Pakistan*.

Stedinger, J. R., Sule, B. F., & Loucks, D. P. (1984). Stochastic dynamic programming model for reservoir optimization. *Water Resources Research*, *20*(11), 1499–1505.

Sunantara, J. D., & Ramirez, J. A. (1997). Optimal stochastic multicrop seasonal and intraseasonal irrigation control. *Journal of Water Resources Planning and Management*, *123*(1), 39–48.

Torabi, M., & Mobasheri, F. (1973). A stochastic dynamic programming model for the optimum operation of the multi purpose reservoir. *Water Resources Bulletin (Harrisburg)*, *9*(6), 1089–1099.

Umamhesh, N. V., & Sreenivasulu, P. (1997). Two-phase stochastic dynamic programming model for optimal operation of irrigation reservoir. *Water Resources Management*, *11*, 395–406.

Vedula, S., & Kumar, D. N. (1996). An integrated model for optimal reservoir operation for irrigation of multiple crops. *Water Resources Research*, *32*(4), 1101–1108.

Vedula, S., & Mujumdar, P. P. (1992). Optimal reservoir operation for irrigation of multiple crops. *Water Resources Research*, *28*(1), 1–9.

Wardlaw, R., & Barnes, J. (1999). Optimal allocation of irrigation water supplies in real time. *Journal of Irrigation and Drainage Engineering*, *125*(6), 345–353.

Warsi, M. (1991). *Indus and Other River Basin of Pakistan, stream flow records*. Case Study Report. WAPDA.

Water And Power Development Authority (WAPDA) (2007). Pakistan.

World Bank (1992). Irrigation planning with environmental consideration. World Bank, Technical Paper 166, Washington, DC.

World Commission on Dams (2000). *Tarbela Dam and related aspects of the Indus River Basin, Pakistan*, A WCD case study research report. Asianics Agro-Dev. International (Pvt) Ltd Cape Town.

Ximing, C., McKinny, D. C., & Lasdon, L. (2003). Integrated hydrologic-agronomic-economic models for river basin management. *Journal of Water Resources Planning and Management*, *129*(1), 4–17.

Yeh, W. W. G. (1985). Reservoir management and operation models: a state-of-the-art review. *Water Resources Research*, *21*(12), 1797–1818.