

RESEARCH ARTICLE

Integrated uncertain models for runoff forecasting and crop planting structure optimization of the Shiyang River Basin, north-west China

Fan ZHANG¹, Mo LI², Shanshan GUO¹, Chenglong ZHANG¹, Ping GUO (✉)¹

¹ Centre for Agricultural Water Research in China, China Agricultural University, Beijing 100083, China
² School of Water Conservancy & Civil Engineering, Northeast Agricultural University, Harbin 150030, China

Abstract To improve the accuracy of runoff forecasting, an uncertain multiple linear regression (UMLR) model is presented in this study. The proposed model avoids the transfer of random error generated in the independent variable to the dependent variable, as this affects prediction accuracy. On this basis, an inexact two-stage stochastic programming (ITSP) model is used for crop planting structure optimization (CPSO) with the inputs that are interval flow values under different probabilities obtained from the UMLR model. The developed system, in which the UMLR model for runoff forecasting and the ITSP model for crop planting structure optimization are integrated, is applied to a real case study. The aim of the developed system is to optimize crops planting area with limited available water resources base on the downstream runoff forecasting in order to obtain the maximum system benefit in the future. The solution obtained can demonstrate the feasibility and suitability of the developed system, and help decision makers to identify reasonable crop planting structure under multiple uncertainties.

Keywords crop planting structure optimization, inexact two-stage stochastic programming, runoff forecasting, Shiyang River Basin, uncertain multiple linear regression

1 Introduction

Increase in water demand and reduction in available water supply are in direct competition, which exacerbates the shortage of agricultural water resources^[1] and this problem is particularly severe in arid and semi-arid areas. To make full use of limited water resources for agricultural production, crop planting structure optimization is an

important approach to increase agricultural economic benefits and improve agricultural water management^[2–4].

CPSO is a complex system with many uncertain parameters, such as crop planting area, irrigation water use efficiency, available water supply and economic parameters. The comprehensive benefits of the system are influenced by all those parameters and phenomena due to their uncertainty characteristics^[5,6]. If defined parameters or models are employed simply instead of uncertain ones, unreliable results will be obtained with important information missing^[7,8]. Among all the uncertain parameters, surface water is quite significant and will directly affect the optimal scheme of crop planting structure adjustment. The main source of the available surface water is catchment runoff. Hence, for accurately optimizing crop planting structure in the future, the accurate prediction of catchment runoff is desirable. Much research has focused on hydrological forecasting methods in order to obtain more accurate predictions for runoff^[9–11]. Of the methods evaluated, a multiple linear regression (MLR) model has proven to be an effective forecast method^[5,12]. In the process of prediction, the random error generated in the prediction of the independent variable can transfer to the dependent variable, which seriously influences the prediction accuracy. In addition, climate change and human activities exacerbate the complex uncertainties in runoff forecasting. To date, few studies have attempted to overcome these disadvantages. Consequently, improvements in the uncertain MLR method are still needed.

Inexact two-stage stochastic programming (ITSP) modes are capable of dealing with the interval parameter with the lower and upper bounds combining with the stochastic process^[13,14]. ITSP models have been widely applied to CPSO owing to the stochastic characteristics in available water resources. For example, ITSP was used to deal with not only the stochastic nature of different water availabilities in different hydrological years, but also the interval uncertainty that emerges in statistics of crop prices

Received March 10, 2017; accepted October 3, 2017

Correspondence: guop@cau.edu.cn

and costs^[15]. However, previous studies have not been able to integrate runoff forecasting under uncertainty with ITSP, and thus cannot provide dynamic decision suggestions for crop planting structure adjustment.

Therefore, this study aimed to improve the MLR model to avoid poor accuracy of the forecast results, and optimize the crop planting structure based on the results of an improved uncertain multiple linear regression (UMLR) model. The improved model was then applied to a case study to evaluate how the system can be used to adjust crop planting structure more efficiently. The study encompassed: (1) the formulation of an UMLR model based on MLR model for runoff forecasting, (2) the development of an ITSP model for CPSO model integrating the flow values under different probabilities, which are the results from the UMLR model, (3) the application of the improved model to a real case study in the Shiyang River Basin, north-west China, and (4) the analysis the results under multiple scenarios to provide recommendations for crop planting structure adjustment.

2 Study system

2.1 Study area

The study area was the Shiyang River Basin (101°41'–104°16' E, 36°29'–39°27' N) (Fig. 1), which is one of three continental rivers in the Hexi corridor of Gansu Province, north-west China. The Shiyang River Basin has a typical arid continental climate, characterized by low and irregular rainfall, high evaporation and drought periods. It has the most serious water shortage in the Hexi Corridor^[16–18]. The Shiyang River originates in the Qilian Mountain and consists of eight tributaries, Dajing, Gulang, Huangyang, Zamu, Jinta, Xiyang, Dongda and Xida rivers (east to

west). All rivers except the Dajing and converge at the Shiyang River and then flow into the Minqin Basin^[19]. Minqin County is located in the lower reaches of the Shiyang River Basin, with an area of 41400 km², which is one of the most important agricultural counties of Wuwei. Through the Caiqi Hydrologic Station, the Shiyang River flows into the Hongya Mountain Reservoir, the largest desert reservoir in Asia. The available water resources in Minqin County are mainly supplied by surface water from the Caiqi Station, groundwater and diverted water. About 70% to 80% of the water consumption is used for agricultural irrigation, which is the largest water user of the whole system. Due to climatic change, population growth, agricultural and economic development stream flow into Minqin County has decreased sharply in recent years. The annual inflow through the Caiqi Station dropped from about 5.78×10^8 m³ in 1955 to less than 1.00×10^8 m³ in 2004.

Poor management practices and unsuitable crop land use make the conflict between water supply and water demand more intense. To guarantee sustainable development of the region, the regional development planning requires that water allocation must fulfill domestic and basic ecological water demand first, then industrial water, which means adjustment of agricultural water use is imperative. Therefore, appropriate agricultural planning within the available water resources, taking into account changing stream flows, is crucial not only for agricultural production but also socio-economic development and ecological restoration in the Minqin Basin. It is important to accurately predict runoff from the Shiyang River into the Minqin Basin, as a guide for the adjustment of crop planting structure.

2.2 Data collection

Meteorological data for the Wuwei Sub-basin, streamflow



Fig. 1 Location of Shiyang River Basin in north-west China

data for the Shiyang River Basin, and some basic regional data for the Wuwei Sub-basin were used in this study. Meteorological data for the Wuwei Sub-basin including annual mean temperature and annual mean precipitation data, were obtained from the China Meteorological Data Sharing Service System. Flow data for the six tributaries obtained from their respective observation stations, and inflow data for Minqin Country were from the Caiqi Station. Population of Wuwei Sub-basin, effective irrigated area and the value of agricultural output, were obtained from Statistical Yearbook of Wuwei City^[20]. All the data above were for the period 1955 to 2014.

Moreover, to predict the appropriate crop planting structure for Minqin Basin under dynamic inflow, the data used in the optimization model were as follows. Area data and crop output per unit area for 2009 to 2014 were from Statistical Yearbook of Wuwei City^[20]. The demand for food and vegetables from Dietary Guidelines for Chinese^[21]. Groundwater and water diversion data were collected from the Shiyang River Basin Key Governance Projects^[22].

2.3 System framework

This paper attempts to combine the optimization model for crop planting structure with runoff forecasting, where the runoff forecasting uses the improved UMLR model to enhance the accuracy of prediction results, and the optimization part applies the ITSP model to optimize crop planting structure. The common point between runoff forecasting and optimization is the available surface water of Minqin Country. A brief description is given in Fig. 2.

3 UMLR model for runoff forecasting

3.1 UMLR model

The MLR model, one of the earliest models used in medium and long-term hydrological forecasting, is still widely applied to related research^[12,23]. The MLR model is usually given mathematically as:

$$R = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e \quad (1)$$

where the dependent variable R represents the reconstructed/forecasting annual downstream flow; the independent variables X_1, X_2, \dots, X_n denote the factors that affect runoff; The constants $\beta_1, \beta_2, \dots, \beta_n$ are the regression coefficients obtained by the least-squares method; and parameter e is the random error with a normal distribution.

However, when the MLR model is used to predict runoff, the values of multiple independent variables are difficult to predict accurately. Consequently, the propagation of random errors will be passed to the predicted value through the MLR model, which will obviously reduce prediction accuracy. To improve the prediction accuracy of runoff forecasting and make full use of uncertain information in independent variables, an improved UMLR model was developed in this study. This UMLR model decomposes the independent variables in the MLR model into the sum of the change trend and the random error. It is described as

$$X_n = x_n + \alpha_n \quad (2)$$

where x_n is the trend of the independent variable; and α_n is the random error of the true value deviates from the trend

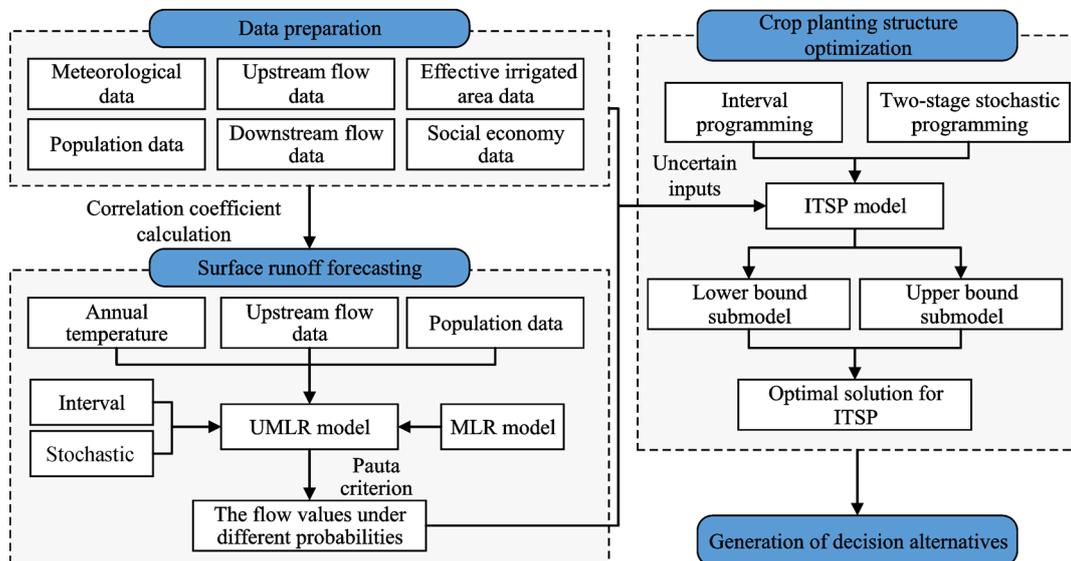


Fig. 2 System framework of integrated uncertain models in this research

with a normal distribution. By combining Eq. (1) and Eq. (2), Eq. (3) is obtained. This integrates the random error from the MLR model and decomposition.

$$R = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + (\alpha_1 \beta_1 + \alpha_2 \beta_2 + \dots + \beta_n \alpha_n + e) \quad (3)$$

Finally, UMLR model is formulated as:

$$R = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \mu \quad (4)$$

where parameter μ is the sum of random errors from the MLR model and decomposition with a normal distribution.

3.2 Runoff prediction

According to previous studies^[5,24,25], the change in runoff from upstream and downstream of the Shiyang River Basin occurred in about 1975. Therefore, the runoff series of the Shiyang River Basin from 1975 to 2014 was selected to ensure a stable sequence. To ensure the regression model effectiveness, the independent variables should be chosen carefully. The criteria for the selection of independent variables are: (1) the independent variables should have a significant effect on the dependent variable with a close linear correlation; (2) the linear correlation between the independent variable and the dependent variable must be authentic not just formal; (3) there should be a certain degree of mutual exclusion between the independent variables, that is, the degree of correlation among the independent variables should not be higher than the degree of correlation with the dependent variable; and (4) the independent variables should have complete statistical data with predictions that are easily to determined.

The correlation coefficients of variables that may affect the inflow in the Caiqi Station are shown in Table 1. In

accordance with the criteria for selection of independent variables, the total flow of six tributaries at mountain outlets, the population of Wuwei Sub-basin, and annual mean temperature were selected as the independent variables of the MLR model to simulate and predict the inflow in the Caiqi Station. The regression equation is:

$$R = 6.530 + 0.095 U - 0.887 P - 0.837 T + e \quad (5)$$

where R represents the inflow in the Caiqi Station; U is the total flow of six tributaries at mountain outlets; P denotes the population of Wuwei Sub-basin; T means the annual mean temperature of Wuwei Sub-basin; e is the random residuals with normal distribution.

To assess the accuracy of the forecasts, deterministic coefficient (DC) and qualified rate (QR) were used^[26]. The values of these two indicators were $DC = 0.91$ and $QR = 75.0\%$, respectively, demonstrating that the precision of the forecast is relatively reliable in accordance with the normal standards of the People's Republic of China, Standard for Hydrological Information and Hydrological Forecasting^[21] for accuracy of runoff forecasting.

The linear model was used to simulate the total flow of six of the tributaries at their mountain outlets and the annual mean temperature of Wuwei Sub-basin from 1955 to 2014, and to forecast trends in these variables from 2015 to 2020 (Fig. 3). Figure 3 shows that the annual total flow of six tributaries at their mountain outlets showed an obviously declining trend. In contrast, the annual mean temperature had a rising trend. Through the linear model, the actual value between 1955 and 2014 was partitioned into the trend and the random residual.

The exponential model was used to forecast the natural growth rate for the population of Wuwei Sub-basin in Shiyang River Basin north-west China from 2015 to 2020. The trend and the random residual are shown in Fig. 4.

Table 1 Correlated coefficients of variable affecting inflow at Caiqi Hydrology Station in the Shiyang River Basin north-west China

Variables	Total flow of six tributaries at mountain outlets	Population of Wuwei Sub-basin	Effective irrigated area	Annual mean temperature	Annual mean precipitation	The price of agricultural total output value	Inflow in the Caiqi Station
Total flow of six tributaries at mountain outlets	1.00	-	-	-	-	-	-
Population of Wuwei sub-basin	-0.22	1.00	-	-	-	-	-
Effective irrigated area	-0.20	0.89	1.00	-	-	-	-
Annual mean temperature	-0.28	0.75	0.62	1.00	-	-	-
Annual mean precipitation	0.28	0.18	0.24	-0.17	1.00	-	-
The price of agricultural total output value	-0.21	0.61	0.46	0.73	-0.14	1.00	-
Inflow in the Caiqi Station	0.45	-0.88	-0.73	-0.84	0.02	-0.75	1.00

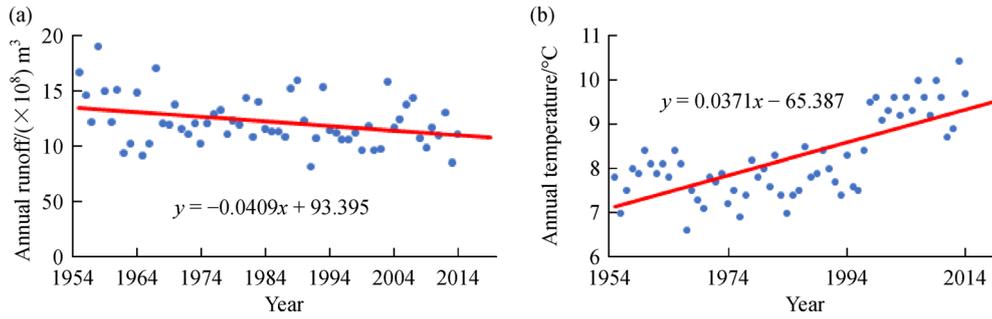


Fig. 3 The total flow of six tributaries at mountain outlets (a) and the annual mean temperature (b) of Wuwei Sub-basin in Shiyang River Basin north-west China simulation and forecast with linear model

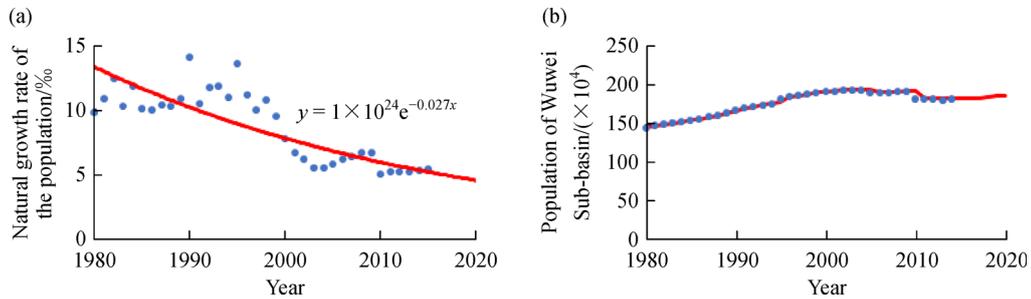


Fig. 4 The natural growth rate of the population (a) and the population of Wuwei Sub-basin in Shiyang River Basin north-west China forecasting result (b)

Using the above values in the UMLR model, the trend of inflow in the Caiqi Station in the period 1975 to 2020 was calculated. At the same time, random error μ was fitted by normal distribution (Fig. 5). The UMLR model avoids the metastasis of the random error from the independent variable prediction to the inflow in the Caiqi Station prediction, and provide not only possible intervals but also the probability of each interval. To obtain the flow values under different probabilities at the Caiqi Station from 2015 to 2020, the fitted normal distribution was divided into three discrete probability intervals. The Pauta criterion, a method distinguishing rough error by judging the calculated probability of interval beyond the interval of 3σ , was used to obtain probability interval boundaries. Then, the normal distribution was divided at probabilities of 75% (low flow level), 50% (medium flow level), and 25% (high

flow level), and three probability intervals of inflow in the Caiqi Station from 2015 to 2020 were obtained (Fig. 6).

Figure 6 shows that the inflow in the Caiqi Station is predicted to have a downward trend from 2015 to 2020. This indicates the necessity to optimize crop planting structure because water available to agriculture will reduce in the future. This forecast estimates the f surface water available in the future for CPSO.

4 ITSP model for crop planting structure optimization

4.1 ITSP model

Two-stage stochastic programming (TSP) refers to a

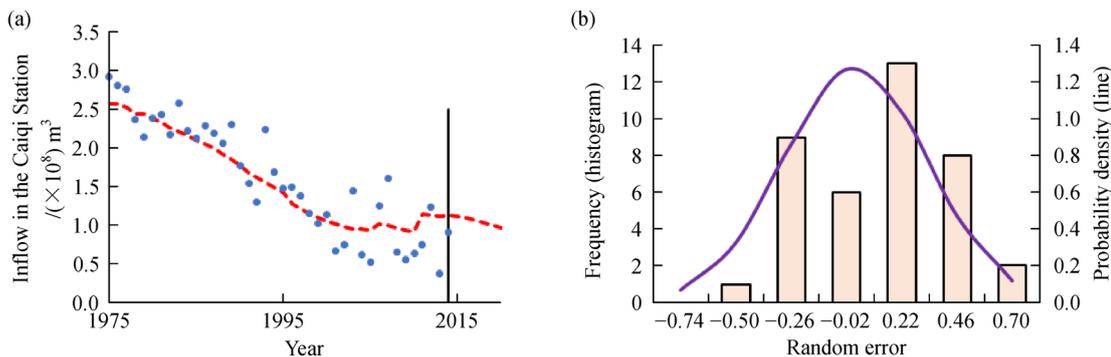


Fig. 5 The trend of inflow at the Caiqi Station, Shiyang River Basin north-west of China (a) and normal distribution fitting of μ (b)

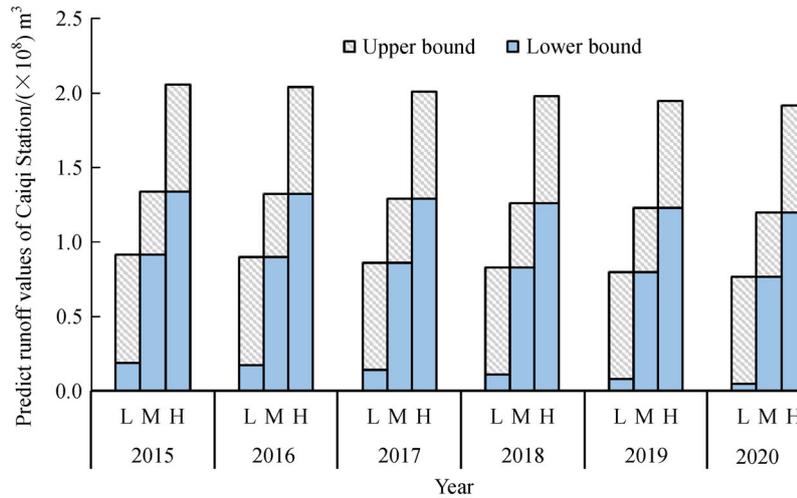


Fig. 6 Three probability intervals of inflow in the Caiqi Station, Shiyang River Basin north-west of China from 2015 to 2020

tradeoff between predefined strategies and the associated adaptive adjustments^[27]. The TSP model can correct the decision in the first stage through making the decision in the second stage after the random event happened in order to reduce the decision risk and obtain the best compromise schemes^[28,29]. In addition, owing to the uncertainties, most real-world problems may lead to the introduction of ambiguous inexact parameters into the TSP framework to result in an ITSP model^[13,30].

ITSP was used to solve many practical problems^[31–33]. The ITSP model for crop planting structural optimization with a maximized planting benefit objective can be mathematically expressed as^[15]:

$$\text{Max } f^\pm = \sum_{i=1}^I A_i^\pm R_i^\pm - \sum_{i=1}^I \sum_{k=1}^K p_k L_i^\pm A_{ik}^\pm \quad \forall i, k \quad (6a)$$

where the subscripts $i = 1, 2, \dots, I$ are the types of crops and I is the total number of crop types, $I = 5$ with $i = 1, 2, 3$ representing grain crops and $i = 4, 5$ representing economic crops; $k = 1, 2, \dots, K$ represents various scenarios of available water levels and K is the total number of scenarios of available water level, with $k = 1$ representing the low flow level, $k = 2$ representing the medium flow level and $k = 3$ representing the high flow level. f (in a currency unit, CNY) denotes the desired value of planting benefit; R_i (CNY · hm⁻²) is the net income per unit area of crop i ; p_k is the probability of occurrence of level k ; L_i (CNY · hm⁻²) is the loss of crop i in the absence of sufficient water, and $L_i > R_i$; the decision variable A_i and A'_{ik} (hm²) are the first stage decision crop area and second stage decision crop area, respectively; and the superscripts “+” and “-” are the upper and lower bounds of the corresponding parameters, respectively (e.g., the upper and lower bounds of the decision variable A_i^\pm can be expressed as A_i^+ and A_i^-).

Subject to:

$$\sum_{i=1}^I M_i (A_i^\pm - A'_{ik}^\pm) \leq Q_k^\pm \eta \quad \forall i, k \quad (6b)$$

(available water constraints)

$$a \leq \frac{\sum_{i=1}^2 (A_i^\pm - A'_{ik}^\pm)}{5} \leq b \quad \forall i, k \quad (6c)$$

(regional crop structure planning constraints)

$$(A_i^\pm - A'_{ik}^\pm) Y_i \geq D_i \quad \forall i, k \quad (6d)$$

(food security constraints)

$$A_i^\pm \geq A'_{ik}^\pm \geq 0 \quad \forall i, k \quad (6e)$$

(non-negativity constraints)

where η is the irrigation water utilization efficiency; Q_k^\pm (10⁸ m³) represents the total amount of available water of selected crops in scene k , including surface water, ground water and water transfer from other regions; M_i (m³ · hm⁻²) is the irrigation quota of crop i ; a and b are the ratio of grain crops planting area and economic crop area, which should satisfy the general plan of the region; Y_i (kg · hm⁻²) is the yield of crop i ; and D_i (kg) is the minimum demand of crop i in this region.

Since the objective is to maximize the system benefits, the sub-model corresponding to f^+ should be calculated first. According to the related researches^[28,33,34], this model (Eqs. (6a)–(6e)) can be solved by decomposing it into two sub-model as follows,

The upper bound value of objective function f^+ :

$$\text{Max } f^+ = \sum_{i=1}^I (A_i^- + \Delta A_i z_i) R_i^+ - \sum_{i=1}^I \sum_{k=1}^K p_k L_i^- A_{ik}^- \quad \forall i, k \quad (7a)$$

subject to:

$$\sum_{i=1}^I M_i [(A_i^- + \Delta A_i z_i) - A_{ik}^{\prime-}] \leq Q_k^+ \eta \quad \forall i, k \quad (7b)$$

$$a \leq \frac{\sum_{i=1}^2 [(A_i^- + \Delta A_i z_i) - A_{ik}^{\prime-}]}{5} \leq b \quad \forall i, k \quad (7c)$$

$$\sum_{i=3}^2 [(A_i^- + \Delta A_i z_i) - A_{ik}^{\prime-}]$$

$$[(A_i^- + \Delta A_i z_i) - A_{ik}^{\prime-}] Y_i \geq D_i \quad \forall i, k \quad (7d)$$

$$A_i^- + \Delta A_i z_i \geq A_{ik}^{\prime-} \geq 0 \quad \forall i, k \quad (7e)$$

$$0 \leq z_i \leq 1 \quad (7f)$$

where z_i and $A_{ik}^{\prime-}$ are decision variables. $z_{i \text{ opt}}$, $A_{ik \text{ opt}}^{\prime-}$, and f_{opt}^+ are the solutions of the sub-model (Eqs. (7a)–(7f)). The optimized first-stage variable can be obtained by $A_{i \text{ opt}}^{\pm} = A_i^- + A_i \cdot z_{i \text{ opt}}$, which may correspond to optimized upper bound objective function value.

The lower bound value of objective function f^- :

$$\text{Max } f^- = \sum_{i=1}^I (A_i^- + \Delta A_i z_{i \text{ opt}}) R_i^- - \sum_{i=1}^I \sum_{k=1}^K p_k L_i^+ A_{ik}^{\prime+} \quad \forall i, k \quad (8a)$$

subject to:

$$\sum_{i=1}^I M_i [(A_i^- + \Delta A_i z_{i \text{ opt}}) - A_{ik}^{\prime+}] \leq Q_k^- \cdot \eta \quad \forall i, k \quad (8b)$$

$$a \leq \frac{\sum_{i=1}^2 [(A_i^- + \Delta A_i z_{i \text{ opt}}) - A_{ik}^{\prime+}]}{5} \leq b \quad \forall i, k \quad (8c)$$

$$\sum_{i=3}^2 [(A_i^- + \Delta A_i z_{i \text{ opt}}) - A_{ik}^{\prime+}]$$

$$[(A_i^- + \Delta A_i z_{i \text{ opt}}) - A_{ik}^{\prime+}] Y_i \geq D_i \quad \forall i, k \quad (8d)$$

$$A_i^- + \Delta A_i z_i \geq A_{ik}^{\prime+} \geq A_{ik \text{ opt}}^{\prime-} \quad \forall i, k \quad (8e)$$

where $A_{ik}^{\prime+}$ is the decision variable. $A_{ik \text{ opt}}^{\prime+}$ and f_{opt}^- are solutions of the sub-model (Eqs. (8a)–(8e)). Thus, optimal solutions of model (Eqs. (6a)–(6e)) can be obtained as:

$$f_{\text{opt}}^{\pm} = [f_{\text{opt}}^-, f_{\text{opt}}^+] \quad (9)$$

$$A_{i \text{ opt}}^{\pm} = A_i^- + \Delta A_i \cdot z_{i \text{ opt}} \quad (10)$$

$$A_{ik \text{ opt}}^{\pm} = [A_{ik \text{ opt}}^{\prime-}, A_{ik \text{ opt}}^{\prime+}] \quad (11)$$

Optimized area of selected crop can be give as:

$$[(A_i^- + \Delta A_i z_{i \text{ opt}}) - A_{ik}^{\prime+}, (A_i^- + \Delta A_i z_{i \text{ opt}}) - A_{ik}^{\prime-}] \quad (12)$$

4.2 Analysis and discussion

The inflow in the Hongya Mountain Reservoir includes the water upstream of the Shiyang River and diverted water from other regions. The diverted water ranges from about $(1.99\text{--}2.33) \times 10^8 \text{ m}^3 \cdot \text{yr}^{-1}$ from Liangzhou District and Jingdian Reservoir. The maximum recoverable amount of groundwater is $1.16 \times 10^8 \text{ m}^3$. According to The Shiyang River Basin Key Governance Projects^[22], the ratio of water available to agriculture and the total water available in Minqin County from 2015 to 2020 can be obtained as 0.56, 0.62, 0.61, 0.61, 0.61, and 0.60. Wheat, corn, vegetables, sunflower and cotton are the five major crops in Minqin County accounting for 80% of the total area planted, so 80% of the water available to agriculture was taken as the available water for the planting structure optimization for these five crops (Table 2). Other data for the five crops are listed in Table 3. These data were inputs of the ITSP model for CPSO.

The submodels (Eqs. (7a)–(7e) and Eqs. (8a)–(8e)) of the ITSP model were solved under the limited available water of selected crops, then the optimized results for each selected crop compared with prevailing conditions (crop planting area in 2014) (Fig. 7).

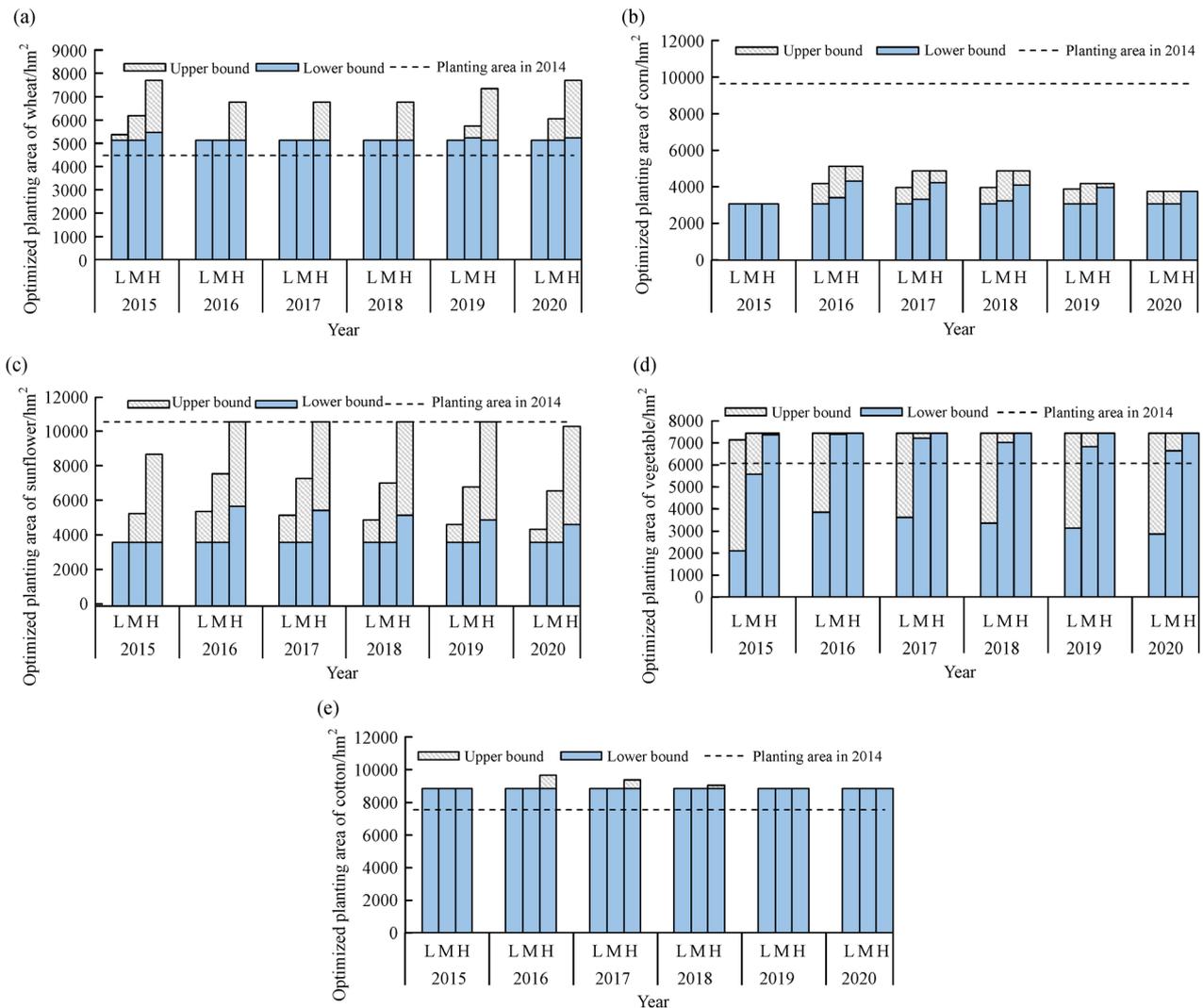
Figure 7 shows that the inflow at Caiqi Station from 2015 to 2020 obtained by UMLR model based on available water predictions affects the optimal planting area of the

Table 2 Range ([upper bound, lower bound] $\times 10^8 \text{ m}^3 \cdot \text{yr}^{-1}$) of available water for the five major crops in Minqin County, Shiyang River Basin north-west of China

Year	Low flow level (75%)	Medium flow level (50%)	High flow level (25%)
2015	[1.35, 1.76]	[1.62, 1.92]	[1.78, 2.20]
2016	[1.49, 1.95]	[1.79, 2.13]	[1.97, 2.43]
2017	[1.47, 1.93]	[1.77, 2.11]	[1.95, 2.41]
2018	[1.45, 1.91]	[1.75, 2.08]	[1.93, 2.39]
2019	[1.43, 1.88]	[1.73, 2.06]	[1.91, 2.36]
2020	[1.41, 1.86]	[1.71, 2.04]	[1.88, 2.34]

Table 3 Associated data for five major crops in the Minqin County, Shiyang River Basin north-west of China (data in square brackets are lower and upper bounds)

Target	Unit	Grain crop		Economic crop		
		Wheat	Corn	Sunflower	Vegetables	Cotton
Planting area	hm ²	4746.67	9853.33	11513.33	6373.33	5000.00
Yield per unit	kg·hm ⁻²	7251.91	10322.61	5566.05	59184.26	1800.02
Price per unit	CNY·kg ⁻¹	[1.80, 2.60]	[2.20, 2.66]	[6.50, 12.00]	[1.80, 2.90]	[15.12, 20.30]
Cost	CNY·hm ⁻²	5850.00	9250.00	3598.00	21547.00	11340.00
Irrigation quota	m ³ ·hm ⁻²	3900.00	4500.00	3450.00	5100.00	3300.00
Net income per unit area	CNY·hm ⁻²	[7203.44, 13004.97]	[13459.75, 18208.15]	[32581.35, 63194.65]	[84984.66, 150087.35]	[15876.36, 25200.49]
Penalty coefficient	CNY·hm ⁻²	[13297.19, 19098.72]	[20491.00, 25239.40]	[34737.60, 65350.90]	[88172.16, 153274.85]	[17938.86, 27262.99]
Minimum demand for crops	t	37049.09	31500.00	20520.00	45990.00	15847.16

**Fig. 7** Optimized planting area of five main crops in Minqin County, Shiyang River Basin north-west of China. (a) Wheat; (b) corn; (c) vegetables; (d) sunflower; (e) cotton.

five main crops. With the goal of guaranteeing food security, the optimal results tend to allocate less land to the two grain crops as these provide lower economic benefits. The area planted to wheat should be increased and the area planted to corn reduced depending on the prevailing conditions. Likewise, the area planted to cotton should be increased appropriately based on the actual condition to meet the demand in the region, and the sunflower area should be substantially reduced compared to the actual condition because of its higher water consumption and lower economic returns. In low flow years, it would be appropriate to reduce the area planted to vegetables, as this can significantly improve the overall economic benefit with plenty of water. These results provide decision makers the possible available water amounts and a wide variety of support for decisions about planting area with different amounts of available water to ensure the supply of basic agricultural products in the region.

Figure 8 shows the comparison of the optimized planting area with the prevailing conditions for the total of the five selected crops at different flow levels. It is clear that the total optimized planting area is less than the existing area, that is, the total optimized planting area is difficult to reach or exceed the existing area, which effectively meets the requirements of The Shiyang River

Basin Key Governance Projects^[22]. The mathematical expectation of the economic benefits, i.e., the objective of the ITSP model, is shown in Fig. 9. Compared to the average total benefits of the five crops planted in 2014, the optimized planting income clearly increases, which successfully avoids the risk of low benefits.

For policy makers in Minqin County, broad decision support is provided for various crop planting scenarios. The optimal planting area of each crop in three possible available water scenarios is provided in Fig. 7. Optimistic policy makers prefer the planting area in high flow level with higher risk of water shortage, and moderate policy makers prefer medium flow level with medium risk, while pessimistic policy makers prefer low flow level with lower risk. The optimal crop area at each flow level is also given as the interval form, which brings greater decision-making options for policy makers. To guarantee enough grain supply in the future, especially corn, policy makers should attempt to provide subsidies or raise the purchase price of grain crops because of their low economic benefits. If there is enough water for irrigating crops, vegetables should be a priority because of their shorter growth period and higher benefits. However, the dynamics of market prices cannot be ignored when formulating the crop planting policy. With the above measures, water shortage problem can be

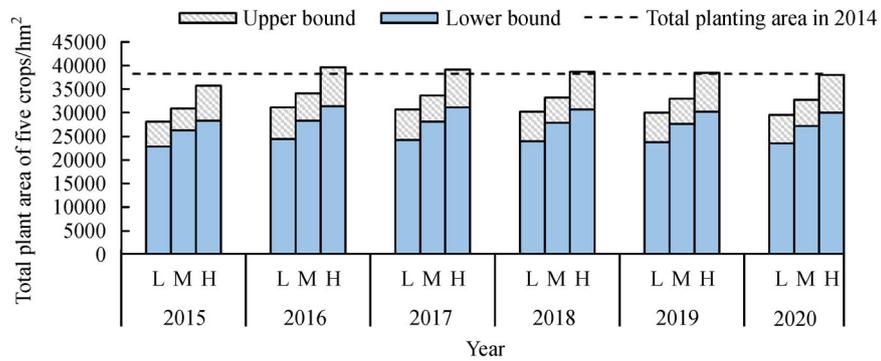


Fig. 8 Optimized total planting area of five crops in Minqin County, Shiyang River Basin north-west of China

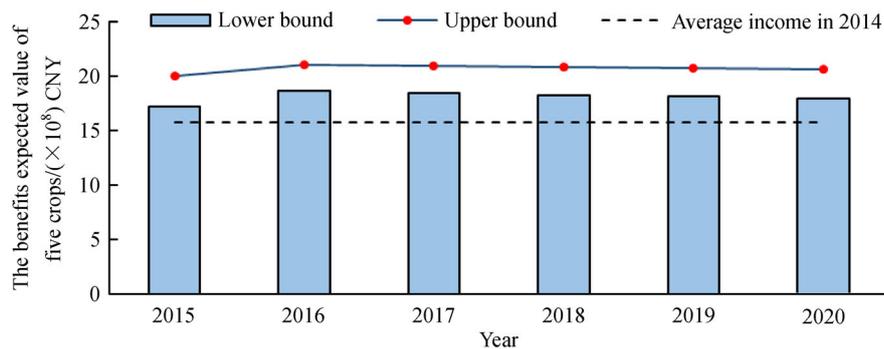


Fig. 9 The total economic benefits mathematical expectation of five crops each year from 2015 to 2020 in Minqin County, Shiyang River Basin north-west of China

adequately adapted to stochastic changes of water availability.

Over all, these analyses indicate that the system, in which the UMLR model and ITSP model are integrated, not only avoids the poor prediction accuracy of the runoff because of the uncertainties in the independent variable prediction, but also simultaneously handles multiple uncertainties of CPSO. This developed system can also be applied to other practical problems related to water resources management.

5 Conclusions

An UMLR model for runoff forecasting was developed. UMLR has the advantage of (1) avoiding the transfer of random error from independent to dependent variables, which affects the prediction accuracy, (2) providing probability for each possible water amounts. Based on the results of the UMLR, the ITSP model is used for CPSO, and it can (1) coordinate the risk and benefits in a framework, (2) address uncertainties expressed as intervals and randomness simultaneously, and (3) provide optimal-planting structures under different amounts of available water.

A case study of CPSO was provided to demonstrate the applicability of the proposed system combining the UMLR model for runoff forecasting with the ITSP model for CPSO. The case study was to optimally allocate crop planting area for selected crops with limited water resources (base on the downstream runoff forecasting) in order to obtain the maximum system net revenue. The comparisons show that the system developed is effective and suitable for arid regions.

This study attempted to improve the MLR model by introducing interval and stochastic uncertainties, and thus an UMLR model was developed. The predicted results of the UMLR model were integrated as the inputs to an ITSP model for CPSO. This system, which combines runoff forecasting and CPSO, can also be applied to other similar regions where the surface water is one of the major sources of water availability, helping decision makers allocate water resources more effectively. However, the UMLR model can hardly predict long-term stream flow because of the shortcomings of the MLR model. Fuzzy methods dealing with uncertainties and other complex methods have not been fully considered in CPSO but it would be desirable to include them in further research.

Acknowledgements This research was supported by the National Key Research and Development Plan of China (2016YFC0400207), the National Natural Science Foundation of China (51439006), and the National High Technology Research and Development Program of China (2013AA102904).

Compliance with ethics guidelines Fan Zhang, Mo Li, Shanshan Guo, Chenglong Zhang, and Ping Guo declare that they have no conflicts of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Su X L, Li J F, Singh V P. Optimal allocation of agricultural water resources based on virtual water subdivision in Shiyang River Basin. *Water Resources Management*, 2014, **28**(8): 2243–2257
2. Zeng X T, Kang S Z, Li F S, Zhang L, Guo P. Fuzzy multi-objective linear programming applying to crop area planning. *Agricultural Water Management*, 2010, **98**(1): 134–142
3. Li M, Guo P, Ren C F. Water resources management models based on two-level linear fractional programming method under uncertainty. *Journal of Water Resources Planning and Management*, 2015, **141**(9): 05015001
4. Li M, Guo P, Zhang L D, Zhang C L. Uncertain and multi-objective programming models for crop planting structure optimization. *Frontiers of Agricultural Science and Engineering*, 2016, **3**(1): 34
5. Gui Z Y, Li M, Guo P. Simulation-based inexact fuzzy semi-infinite programming method for agricultural cultivated area planning in the Shiyang River Basin. *Journal of Irrigation and Drainage Engineering*, 2016, **2**: 05016011
6. Tan Q, Zhang S, Li R. Optimal use of agricultural water and land resources through reconfiguring crop planting structure under socioeconomic and ecological objectives. *Water*, 2017, **9**(7): 488
7. Li M, Guo P. A multi-objective optimal allocation model for irrigation water resources under multiple uncertainties. *Applied Mathematical Modelling*, 2014, **38**(19–20): 4897–4911
8. Zhang C L, Guo P. A generalized fuzzy credibility-constrained linear fractional programming approach for optimal irrigation water allocation under uncertainty. *Journal of Hydrology*, 2017, **553**: 735–749
9. Trivedi H V, Singh J K. Application of grey system theory in the development of a runoff prediction model. *Biosystems Engineering*, 2005, **92**(4): 521–526
10. Maier H R, Dandy G C. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. *Environmental Modelling & Software*, 2000, **15**(1): 101–124
11. Guo J, Zhou J Z, Qin H, Zou Q, Li Q. Monthly streamflow forecasting based on improved support vector machine model. *Expert Systems with Applications*, 2011, **38**(10): 13073–13081
12. Driver N E, Troutman B M. Regression models for estimating urban storm-runoff quality and quantity in the United States. *Journal of Hydrology*, 1989, **109**(3): 221–236
13. Huang G H, Loucks D P. An inexact two-stage stochastic programming model for water resources management under uncertainty. *Civil Engineering and Environmental Systems*, 2000, **17**(2): 95–118
14. Zhang C L, Guo P. An inexact CVaR two-stage mixed-integer linear programming approach for agricultural water management under uncertainty considering ecological water requirement. *Ecological Indicators*, 2017,
15. Zhang F, Guo P, Li M. Planting structure optimization of main crops in the middle reaches of Heihe River Basin based on dual interval

- two stage stochastic programming. *Journal of China Agricultural University*, 2016, (11): 109–116 (in Chinese)
16. Shi Q, Chen F H, Zhu Y, Madsen D. Lake evolution of the terminal area of Shiyang River drainage in arid China since the last glaciation. *Quaternary International*, 2002, **93**(3): 31–43
 17. Wang Q, Shi J, Chen G, Xue L. Environmental effects induced by human activities in arid Shiyang River Basin, Gansu Province, northwest China. *Environmental Geology*, 2002, **43**(1–2): 219–227
 18. Kang S Z, Su X L, Tong L, Du T S. The impacts of human activities on the water–land environment of the Shiyang River Basin, an arid region in northwest China. *Hydrological Sciences Journal*, 2009, **49**(3): 427
 19. Zhang L D, Guo P, Fang S Q, Li M. Monthly optimal reservoirs operation for multicrop deficit irrigation under fuzzy stochastic uncertainties. *Journal of Applied Mathematics*, 2014, **2014**(2): 1–11
 20. Statistics Bureau of Wuwei. Statistical Yearbook of Wuwei City. *Wuwei*, 2009–2014 (in Chinese)
 21. National Health and Family Planning Commission. Dietary Guidelines for Chinese. Beijing: *People's Publishing House*, 2016 (in Chinese)
 22. Gansu Provincial Department of Water Resources. Shiyang River Basin Key Governance Projects. *Wuwei*, 2007 (in Chinese)
 23. Min F, Wu X D. Local Semi-linear Regression for River Runoff Forecasting: IEEE Cse' 09, *IEEE International Conference on Computational Science and Engineering*, 2009, 29–31
 24. Dong P G, Feng S Y, Huo Z L. Effect of climatic changes and human activities on the runoff at down-stream of the Shiyang River. *China Rural Water and Hydropower*, 2010, (7): 22–24 (in Chinese)
 25. Huo Z L, Feng S Y, Kang S Z, Li W, Chen S. Effect of climate changes and water-related human activities on annual stream flows of the Shiyang River Basin in arid north-west China. *Hydrological Processes*, 2010, **22**(16): 3155–3167
 26. AQSIQ (General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China) and SAC (Standardization Administration of the People's Republic of China). *Standard for hydrological information and hydrological forecasting. GB/T 22482–2008*. Beijing, 2009 (in Chinese)
 27. Cui L, Li Y P, Huang G H. Planning an agricultural water resources management system: a two-stage stochastic fractional programming model. *Sustainability*, 2015, **7**(8): 9846–9863
 28. Li Y P, Huang G H, Nie S L, Nie X H, Maqsood I. An interval-parameter two-stage stochastic integer programming model for environmental systems planning under uncertainty. *Engineering Optimization*, 2006, **38**(4): 461–483
 29. Li W, Li Y P, Li C H, Huang G H. An inexact two-stage water management model for planning agricultural irrigation under uncertainty. *Agricultural Water Management*, 2010, **97**(11): 1905–1914
 30. Guo P, Huang G H, He L, Sun B W. ITSSIP: interval-parameter two-stage stochastic semi-infinite programming for environmental management under uncertainty. *Environmental Modelling & Software*, 2008, **23**(12): 1422–1437
 31. Xie Y L, Huang G H, Li W, Li J B, Li Y F. An inexact two-stage stochastic programming model for water resources management in Nansihu Lake Basin, China. *Journal of Environmental Management*, 2013, **127**(2): 188–205
 32. Mo S H, Duan H N, Shen B, Wang D. Interval two-stage stochastic integer programming for urban water resource management under uncertainty. *Journal of Coastal Research*, 2015, **73**: 160–165
 33. Guo P, Huang G H, He L, Zhu H. Interval-parameter two-stage stochastic semi-infinite programming: application to water resources management under uncertainty. *Water Resources Management*, 2009, **23**(5): 1001–1023
 34. Li C Y, Zhang L. An inexact two-stage allocation model for water resources management under uncertainty. *Water Resources Management*, 2015, **29**(6): 1823–1841