Team Control Number

70732

For office use only		For office use only
T1 _____		F1 _____
T2 _____		F2 _____
T3 _____		F3 _____
T4 _____		F4 _____

Problem Chosen

A

**2017
MCM/ICM
Summary Sheet**

Summary

Our goal is to establish a series of models to determine a new dams system along the Zambezi River. After weighing both economic and security factors, we determine the number and location of the dams accordingly. More importantly, we predict the discharge and provide a strategy to handle extreme conditions. We establish four models to achieve this goal.

First, to establish relations between features and benefits as well as costs of dams, we use the Back Propagation(BP) neural network to build the Benefit-Cost Analysis Model. We assume 7 cost-determining indexes and 3 benefit-determining indexes. Data of several worldwide dams incluing 10 indexes above, together with costs and benefits, are obtained from professional institutes. We train the neural network and get 2 results of preferable fits.

Based on the above model, we establish a Decision-Making Model of small dams. Our aim is to determine the number of dams needed, the location of dams as well as the size of each dam. So the number of dams and the position of the most upstream dam are set as two unknowns. For safety concern, we assume the total capacity of small reservoirs equals to Kariba and every small reservoir distributes the same reservoir capacity. Then we build a relationship between unknowns and water levels of all dams. Furthermore, a group of 7 indexes mentioned in the above model are determined according to water level. We make the benefit-cost analysis utilizing those 7 indexes which directly depends on the two unknowns above. After traversing all befitting possibilities of the two unknowns, we manage to find a group of unknowns corresponding to the best economic efficiency. Eventually, we know the number of dams, the position of the most upstream dam and all parameters of every dam.

Then we establish the Flow Design Model to predict the maximum and minimum discharge. In this model, we choose the probability statistic method .We consider the annually discharge Q as a random variable. A mathematical probability distribution, Person-III distribution, is used to express the distribution of discharge variables. We can thus use this distribution to predict the discharge of any frequency, including extreme conditions.

Finally, we build a Connection Reservoirs Operation Model to determine the order of pouring or storing water under extreme conditions. We assume the optimal criterion of operating reservoirs is to achieve maximum electric power. Through rigorous mathematical derivation, we get 2 concise indexes for judging the order of pouring and storing water separately. The 2 indexes are easy to use under extreme conditions such as floods and droughts.

A New Life of Zambezi

Team #70732

January 24, 2017

Contents

1   Introduction

1.1   Background Review

The Kariba Dam built in the 1950s on the Zambezi River is one of the largest dams in Africa. Its construction was controversial, the Institute of Risk Management South Africa Risk Research Report warns people that the dam is in dire need of maintenance. The Zambezi River Authority raises a lot of money and prepares to take some measures, which touches off a wide disscussion of it.

1.2   Restatement of the Problem

We are required to solve two problems:

- Provide a brief assessment of the three options.

- Provide a detailed analysis of Option 3.

  As for the first problem, we calculate costs and benefits of the three options and determine which one is the best.

  As for the second problem, We decompose it into three sub-problems:

  - Determine the number, location and various indexes of the new dams along the Zambezi River. We should weigh both economic and security factors.
  - Predict the maximum/minimum discharge under extreme conditions.
  - Explain and justify the actions that should be taken to properly handle emergency water flow situations (flooding or prolonged low water conditions).

We solve the problem in 3 steps.

Step1. We seek a model to establish relations between features and benefits as well as costs of dams. Based on the above model, we balance security and economic efficiency. Then we determine the number of dams needed, the location of dams as well as the indexes of each dam.

Step2. We use probability statistic model to obtain the discharge of specific frequency. Thus we can predict the maximum and minimum discharge.

Step3. We try to find the best way to operate reservoirs under extreme conditions such as flood or low water.

1.3   Notification

Table 1: Notification

Symbol	Definition	Units	Remarks
i	a corner mark	Unitless	It represents the order of something $i = 1$ represents the first one.
RC	Reservoir Capacity	$10^8 m^3$	Adding the corner mark i means the corresponding property of dam_i
DH	Dam Height	m	Detailed in the Remarks of RC
CV	Concrete Volume	$10^4 m^3$	Detailed in the Remarks of RC
EV	Earth Volume	$10^4 m^3$	Detailed in the Remarks of RC
IC	Installed Capacity	$10^4 kw$	Detailed in the Remarks of RC
CP	Construction Period	$year$	
LEL	Local Economic Level	10^4RMB	
CC	Conversion Cost	10^8RMB	
AGC	Annual Generating Capacity	$10^8 kw \cdot h$	Detailed in the Remarks of RC
CA	Catchment Area	$10^4 km^2$	Detailed in the Remarks of RC
AB	Annual Benefits	10^8RMB	
RC_{Kariba}	Reservoir Capacity of Kariba	$10^8 m^3$	
IC_{Kariba}	Installed Capacity of Kariba	$10^4 kw$	
CV_{Kariba}	Concrete Volume of Kariba	$10^4 m^3$	
CA_{Kariba}	Catchment Area of Kariba	$10^4 km^2$	
PP	Payback Period of investment	$year$	
v	vector	Unitless	
nv	normalized vector	Unitless	
w	the width of all reservoirs	km	
x	the distance between the most upstream small dam and the Zero Point	km	
n	the number of small dams	Unitless	
$f(t)$	Spline function of river profiles	Unitless	
S_i	distance between dam_i and $dam i + 1$	km	
x_i	X-axis of dam_i	km	
h_i	water level of dam_i	m	
l_i	dam width of dam_i	m	
Q	discharge	m^3	
$\Gamma(x)$	a function	Unitless	$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$
C_V	coefficient of variation	Unitless	
C_S	coefficient of skewness	Unitless	

2 A Brief Assessment of the Three Options

2.1 Qualitative Analysis

- **(Option 1)Repairing the existing Kariba Dam**

  - **Costs**

    The costs mainly include the following items:

    a. Materials such as concrete used to repair the dam.

    b. The labor.

    c. Electric energy and water source.

    d. Management and maintenance.

    This Option obviously costs the least money.

  - **Benefits**

    The benefits mainly include the following items

    a. Electricity generation.

    b. Irrigation.

    c. Water supply.

    Repairing the dam prolongs the service life of the dam. Thus, the dam can bring a lot more economic benefits in the long term.

- **(Option 2)Rebuilding the existing Kariba Dam**

  - **Costs**

    The costs of the second option are similar to the first one. But we should take care that rebuilding a dam costs us much more materials and energy than repairing a dam. In addition, there are another two items which we should notice :

    a. Compared with the first option, we need to move the previous dam and build the new dam, which may last over 10 years. Different from the first option, the dam cannot bring any benefit for us during this time, which should be considered as potential cost.

    b. In order to protect downstream areas from flood disaster when removing the previous dam, we need to take some measures, which cost us some money.

  - **Benefits**

    In general, the benefits of the second option are similar to the first one. Compared with 1950s(the Kariba dam was built in 1950s), there are a lot of new technology and materials now. If we rebuild the dam, it will be more powerful and durable. For example, it may provide more electric energy and benefit us more then the first option.

    a. Electricity generation.

    b. Irrigation.

    c. Water supply.

- **(Option 3)Removing the Kariba Dam and replacing it with a series of ten to twenty smaller dams along the Zambezi River.**

  – **Costs**

  In general, the costs of the third option are similar to the second one.

  – **Benefit**

  In general, the benefits of the third option are similar to the second one. But there are 3 another obvious benefits:

    a. A series of dams increase the catchment area affected. Thus, the dams can irrigate and provide water for more areas.

    b. When we need electric energy and water, we choose the energy generated from the nearest dam. The decrease of transportation cost can increase the profit conversely.

    c. Even one dam of the series breaks down, the other dams are in normal state. In other words, the economic effect is sustained.

In a word, each option has its advantages and disadvantages. Thus, we use the economical principles to analyse quantitatively.

2.2 Quantitative Analysis

In order to have a intuitive assessment of the three options, we use some data calculated by our model in Requirement 2, and we find some statistics(such as materials costs). The data are listed as follows:

Table 2: Cost Data of Some Dams

option	costs	CP	benefits				LE	NAV
			Genetating Capacity	Water supply and irrigation	other	sum		
1	0.6	2	0.27	0.19	0.05	0.51	20	0.40
2	5	10	0.49	0.32	0.1	0.31	70	0.14
3	4	5	0.51	0.35	0.12	0.51	100	0.47

1.Costs include materials, labor and energy(billion dollars/year).
2.CP is the short form of construction period(year).
3.The unit of benefits is dollars/year.
4.LE is the short form of life expectancy(year).
5.NAV is the short form of net annual value(billion dollors)

According to the economical principles, we use the data to work out the NAV(a parameter in economics which used to assess the profit of a engineering project)of the three options. We can clearly see that the NAV of option 3 is the largest, which means that the option 3 brings us the best economic profits.

Overall,the option 3 is the best choice for us, and we will analyse the option 3 in detail in the main part of our paper.

3   Benefits-Costs Analysis Model

3.1   Introduction

Back Propagation(BP) is a common method of training artificial neural networks. According to almighty approaching theorem, a 3-layer BP neural network with a hidden layer can approximate to any continuous function of bounded domain with arbitrary precision.[2] 3-layer BP neural network can be explained by figure 1. There are several inputs and outputs, and three layers: the input layer(I), output layer(O) and hide layer(H). The input layer neurons receive the inputs and transfer signals to hidden layer neurons. There is a weight between every two neurons in the near layers. The output layer neurons receive signals from the hidden layer neurons and change them to outputs. By adjusting the weights, they can be very close to the desired outputs.
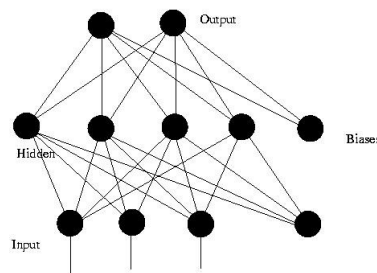
Figure 1: A 3-layer neural network

3.2   Assumptions

Benefits of a dam are various, including generating capacity, flood protection, irrigation, water supply and so on. However, analysing them separately is quite difficult. Professional institutes can give a specific number about the benefit after evaluating the dam, like the annual benefit(AB). But we cannot know how do they get it. Thus, we assume benefits of a dam are mainly determined by 3 indexes listed below and we try to build a network between our indexes and the final benefit.

1. **Reservoir Capacity.** It is obvious in real life that the reservoir capacity determines the scale of the dam, and so do the benefits.

2. **Annual Generating Capacity.** Kariba Dam was mainly built to provide sufficient power for copper mines. Thus we can assume that the benefit is mainly from the annual generating capacity.

3. **Catchment Area.** Catchment area is the whole area that the dam can affect, it usually means the population that people who can benefit from the dam.

Likewise, we assume that the Cost of dam is mainly determined by the following 7 indexes:

1. **Reservoir Capacity.** Reservoir capacity can also determine the costs. The larger reservoir capacity costs more.

2. **Dam Height.** Empirically, dam height is closely related to the scale of the project.

3. **Concrete Volume.** Concrete occupies a large proportion of costs, and it is one of the most visualized index of a dam.

4. **Earth Volume.** Earth volume can also represent the scale of the dam in most cases.

5. **Installed Capacity.** Installed capacity shows how many electric generators are used, and that is a big investment.

6. **Local Economic Level.** This determine the labor cost of the construction project. Generally, we use the GDP per person to serve as the index of economic level.

7. **Construction Period.** Construction project spends enormous sums every day. So the longer construction period is, the higher costs are.

Besides, the cost value is not the past construction cost. Instead, we use Conversion Cost(CC) - The cost today obtained by discounting the rate of income. Set the conversion cost as CC, the past cost as C, the start date(year) as t_0, the current date(year) as t, the rate of income(annual) as r, then:

$$CC = C * (1 + r)^{t - t_0}$$

According to the hydraulic industry standard, the rate of income is 7%. So the above equation changes to the following form:

$$CC = C * (1 + 7\%)^{t - t_0}$$

For the sake of convenience, we still call the conversion cost as cost in the rest of this article.

3.3   Building BP neuron network for Costs

We set the number of the input layer neurons equal to the inputs, so do the output layer neurons. According to assumption(1), there are several indexes closely related to the cost including Reservoir Capacity(RC), Dam Height(DH), Concrete Volume(CV), Earth Volume(EV), Installed Capacity(IC), Construction Period(CP), Local Economic Level (LEL). Then we collect data of several dams to serve as inputs. But these indexes should be normalized before they are used because normalized inputs are much easier to train. A easy and quick normalize algorithm is linear transformation algorithm. For example, set the vector to be normalized as v, the normalized vector as nv, some element of v as v_i, the ith component of nv as nv_i, then:

$$nv_i = \frac{v_i - min(v)}{max(v) - min(v)}$$

These indexes are included in table 7. [3] Then we use Matlab to help to build simulation model for BP neuron network. We use the data of the dams to train the neurons and input them into the model. Since the BP neural network is easy to get into local minimum, we must try many times to get the best result. Figure 2 shows a satisfactory result, the coefficient of correlation(R) reached 0.99883. Once we get the trained neuron network, we can use it to predict the cost with the 7 indexes.

Table 3: Data of Several Dams(1)

Name	LYX	BHT	LWX	LJX	ET	XW	JP	WDD	XLD
$CV/(10^4m^3)$	370	796	258	200	390	1056	474	381	289
DH/m	178	289	250	165	240	292	305	270	295
$RC/(10^8m^3)$	247	206	11	17	58	151	78	76	140
$CP/year$	17	12	6	11	9	8	7	8	14
$LEL/(10^4\text{RMB})$	4.11	3.66	4.11	4.11	3.66	2.87	2.87	2.87	2.87
$IC/(10^4kW)$	128	1600	420	200	330	420	330	1020	1440
$EV/(10^4m^3)$	157	1500	210	160	477	1370	1000	1822	730
$CC/(10^8\text{RMB})$	2246	1270	387	665	1547	715	476	1268	1728
$StartDate/year$	1976	2010	2002	1988	1991	2002	2007	2012	1994
$PastCost$	150	846	150	100	285	277	259	967	390

The names of dams are abbreviated. You can look them up in the appendix.

3.4　Building BP Neuron Network for Benefits

Likewise, we build a BP neuron network for the benefits. What is different from the cost is the input. According to our assumptions, the benefit of the dam is mainly determined by Reservoir Capacity(RC),Annual Generating Capacity(AGC), Catchment Area(CA). The outputs are Annual Benefits(AB). Indexes of these dams are listed in table 4.

Table 4: Data of Several Dams(2)

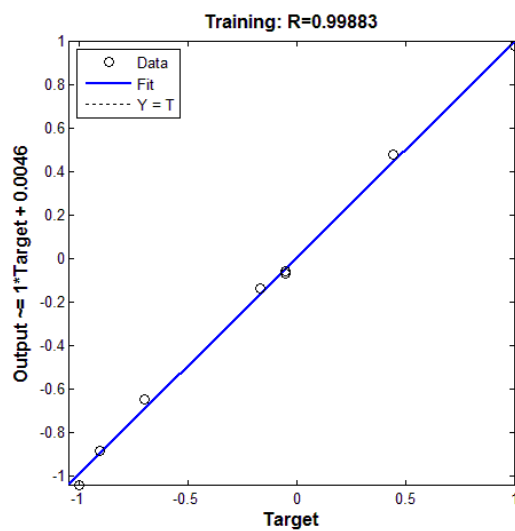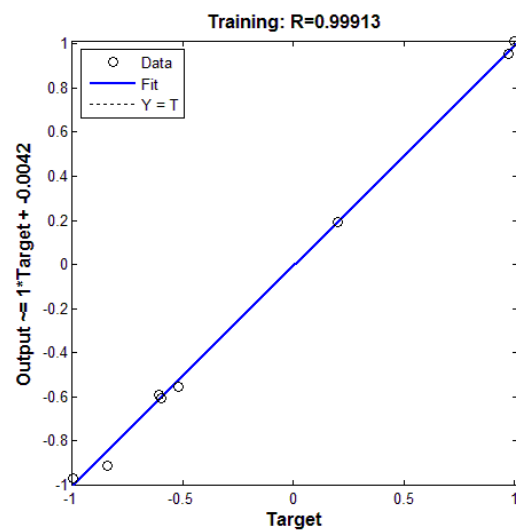| Name	ET	XW	JP	WDD	XLD	LYX	BHT	LWX	LJX
$StartDate/year$	1991.0	2002.0	2007.0	2012.0	1994.0	1976.0	2010.0	2002.0	1988.0
$RC/(10^8m^3)$	58.0	151.3	77.6	76.0	140.2	247.0	206.0	10.6	16.5
$AGC/(10^8kw \cdot h)$	170.4	190.6	166.2	389.3	609.8	59.4	602.4	102.3	59.0
$CA/(10^4km^2)$	11.6	11.3	10.3	40.6	45.4	13.1	43.0	13.2	13.7
$LEL/(10^4\text{RMB})$	3.7	2.9	2.9	2.9	2.9	4.1	3.7	4.1	4.1
$AB/(10^8\text{RMB})$	29.8	33.4	29.1	68.1	106.7	10.4	105.4	17.9	10.3

Figure 2: R of costs

Figure 3: R of benefits

4 Decision-Making Model of Small Dams

4.1 Introduction

With the model above, we get access to predict benefits and costs. In order to evaluate the possibilities and economics of replacing Kariba Dam with several small dams, we establish a Decision-Making Model based on the Benefit-Cost Analysis Model. The problem is that exact number of dams and distance between two dams cannot be determined at present. Thus, we set up to bring forward a best solution on the principle of the smallest Payback Period(PP).

Before beginning, we define the reservoir capacity of Kariba as RC_{Kariba} , and the width of all reservoirs is w, also, it meets that $w = 40km$.

4.2 Assumptions

1. The distance between Kariba and small dams is limited no more than $250km$ and reservoirs originated from the interception of different small dams are independent.

2. We set $40km$ as the width of all reservoirs according to the average width of Kariba Reservoir.

3. The total capacity of all small reservoirs equals to Kariba and every small reservoir distributes the same reservoir capacity.

4. Topographical condition is considered but not the main factor.

5. Benefit-cost analysis is based on the model mentioned above.

6. Dam height is 1.6 times of the water level in view of the extreme conditions.

7. Dam width lies on the position and it is characterized by a step function.

8. The actual distance between adjacent dams is 1.2 times of theoretical distance since an enough safety margin is needed.

9. Reservoir capacity of Kariba $RC_{Kariba} = 180km^3$, and the installed capacity of Kariba $IC_{Kariba} = 1626MW$, and the catchment area of Kariba $CA_{Kariba} = 663000km^2$.

4.3 Individual Benefit-Cost Analysis

1. **Definite the unknowns**

   Given the assumption that the distance between Kariba Dam and small dams cannot exceed $250km$ and the probability that dams distribute upstream or downstream, the extended length from the first dam to the last is set within $500(= 250 + 250)km$. We select one upstream point which is $250km$ distant from Kariba Dam, and name it **Zero Point** . Assume the number of small dams is n, also the distance between the most upstream small dam and the Zero Point is defined as x, and x is no more than $100km$.

2. **Predict the water level with** x, n

- Individual reservoir capacity. We have supposed that every reservoir has the same capacity. Assume the capacity of reservoir i is $RC_i(i = 1, 2, \ldots, n)$, and thus the individual capacity can be expressed as:

$$RC_i = \frac{RC_{Kariba}}{n}$$

- Establish relationships between capacities and water levels of dams since the latter provide a lower limit of dam heights which tie with the eventual cost closely. In addition to the water level, the reservoir capacity hinges on a two-dimensional parameter, water area, which can be transmuted into a one-dimensional condition if we postulate width of the area as a constant value.

- Use river profiles to model the one-dimensional condition. We have collected altitude statistics of river profiles within the confines of $500km$ (shown in Figure 4, and figure 5 shows the profile of $250km$), which contain 200 points in total, and fit all points by spline curves as shown in figure 6. Eventually we get the mathematical expression $f(t)$ of the spline. [5]

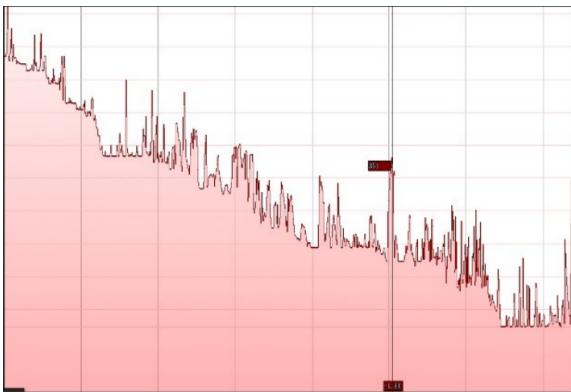Figure 4: the whole range of the dam distribution(in red color)

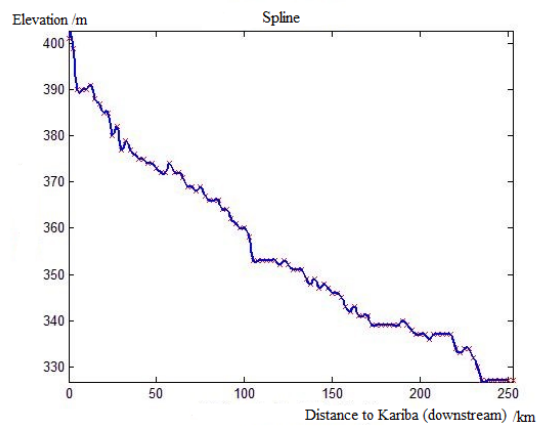Figure 5: primary data of $250km$

Figure 6: spline fitting of $250km$

- Calculate the space between two small dams. Assume S_i is the distance between dam i and dam $i + 1$. Referred to the Zero Point, we define x_i as the X-axis of dam i, the first dam is situated x kilometers away from the Zero Point, so $x1 = x$. Accordingly, x_i can be expressed by x_i and S_i as:

$$x_i = \begin{cases} x + \sum_{k=1}^{i-1} S_k, & i \leq 1 \\ x, & i = 1 \end{cases}$$

Also the reservoir capacity can be expressed as:

$$RC_i = w \cdot \int_{x_{i-1}}^{x_i} (f(x_{i-1}) - f(t))dt$$
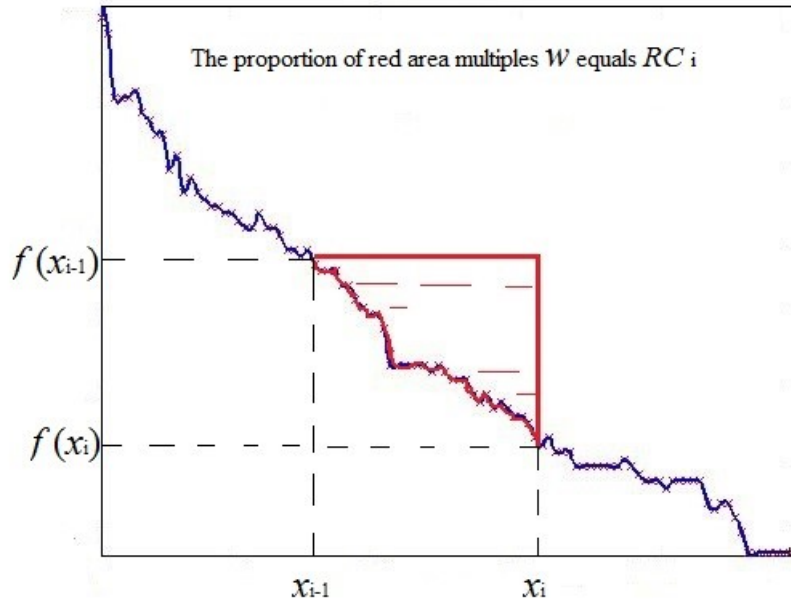
Figure 7: Reservoir Capacity Diagram

Meanwhile, with the precondition that the capacity of a small reservoir equals $\frac{RC_{Kariba}}{n}$, the distance between the first and second dam S_1 should meet:

$$S_1 = \frac{RC_{Kariba}}{n} = w \cdot \int_{x}^{x+S_1} (f(x) - f(t))dt$$

And S_1 can be uniquely determined as long as x is a numerical value. If S_1 is certain, the X-axis of the second dam x_2 can be expressed as:

$$x_2 = x + S_1$$

Repeatedly do above work for n (the number of dams) times and we can obtain the value of $S_i (i = 1, 2, \ldots, n)$ as long as x and n are numerical values, so we cite a functional symbol g_i :

$$S_i = g_i(x, n), i = 1, 2, \ldots, n$$

Furthermore, assume h_i is the water level of dam i , and it meet a condition as follows:

$$h_i = f(x_{i-1}) - f(x_i)$$

3. **parameters prediction with the water level**

- **Height** In view of the extreme conditions, the planning height of dam is considered to be 1.6 times of the water level, so if the height of dam i is represented by DH_i, we get $DH_i = 1.6h_i$.

- **Width** With regard to the dam width l_i of dam i, we assume it is equal to the width of valley, which is characterized by a step function:

$$l_i = \begin{cases} 50m(0km < x_i \leq 125km) \\ 120m(125km < x_i \leq 250km) \\ 200m(250km < x_i \leq 270km) \\ 1000m(270km < x_i \leq 437km) \\ 450m(450km < x_i \leq 500km) \end{cases}$$

- **Concrete Volume** Assume CV_i as the concrete volume of dam i, and there is a positive association between CV_i and $l_i \cdot DH_i$, and a coefficient named α is introduced to estimate the proportion of them. On the basis of several dams in the Model 1, α tend to be set as 10. Consequently, the concrete volume can be expressed as:

$$CV_i = 10 \times DH_i \times l_i = 10 \times 1.6h_i \times l_i = 16h_i l_i$$

- **Earth volume** Define EV_i as the earth volume of dam i , and it is roughly the same as the concrete volume CV_i:

$$EV_i = CV_i = 16h_i l_i$$

- **Installed capacity** Installed capacity depends on the generation benefit which is in close touch with the river fall. In general, the river fall increases along with the water level, and in a premise of ensuring the same benefit or more, the whole installed capacity of all dams cannot be less than Kariba, and we distribute the whole installed capacity to every dam in proportion to the water level. So the installed capacity IC_i of dam i is:

$$IC_i = \frac{h_i}{\sum_{k=1}^{n} h_k} IC_{Kariba}$$

- **Construction Period** Referring to the instances in Model 1, the dam tend to take a long time if the concrete and earth volume are quite enormous. As a result, we set the period to 5 years considering the possibility that several dams start together.

- **Local Economic level** Local Economic level is tightly linked with labor cost. We put to use the GDP per capita of Zambia to evaluate the labor element.

- **Annual Generating Capacity** According to the data in Benefit-Cost Analysis Model, we analyze a fitting function of annual generating capacity AGC_i of dam i with installed capacity IC_i:

$$AGC_i = 0.39IC_i + 5.77 = 0.39\frac{hi}{\sum_{k=1}^{n} h_k} IC_{Kariba} + 5.77$$
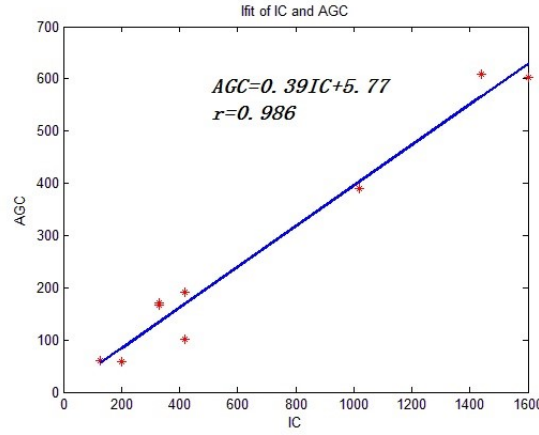
Figure 8: $AGC \sim ICcurve$

- **Catchment Area** Catchment area has something to do with the length of river, and consequently we distribute the whole catchment area of Kariba to every dam in proportion to the distance S_i between two adjacent dams. And the catchment area of reservoir CA_i can be expressed as:

$$CA_i = \frac{g_i(x,n)}{\sum_{k=1}^{n} g_k(x,n)} CA_{Kariba}$$

Totally, every parameter mentioned in the Benefit-Cost Analysis Model is determined by two independent unknowns (x, n), four known quantities $(RC_{Kariba}, CA_{Kariba}, IC_{Kariba})$ and a spline $f(x)$. Note that S_i is changed to $1.2S_i$ for allowance in extreme situation.

4. **Benefit-cost analysis**

Given the preparing work above, one set of data (x, n) correspond to one set of parameters of all dams/reservoirs. The number of small dams n is required to vary in 10 to 20, and the distance between the most upstream small dam and the Zero Point is limited no more than 100 kilometers. Therefore, the number of all befitting possibilities is unlikely to surpass 100*(20-10)=2000 seeing that not all sets of data meet the reservoir capacity requirement in virtue of too small distance of adjacent dams.

According to the Benefit-Cost Analysis Model, both the annual benefit and cost can be predicted through one set of parameters of dam/reservoir. In order to compare with each other objectively, the static payback period of investment named PP is selected to be the central evaluation indicator, and it satisfies the condition as follows:

$$PP = \frac{cost}{annual\ benefit}$$

And we devote ourselves to finding the shortest PP.

4.4 Overall Analysis

1. **Determine x and n of the best situation**

Traverse all sets of data (x, n) and seek out the best one corresponding to the shortest static payback period of investment, and we tend to choose the best set of data as eventual number of dams and distance between the most upstream small dam and the Zero Point. Except for some sets of data missing the requirement, all available are collected to draw a three-dimensional image as below, and the lowest point in the image is the best set of data mentioned above:
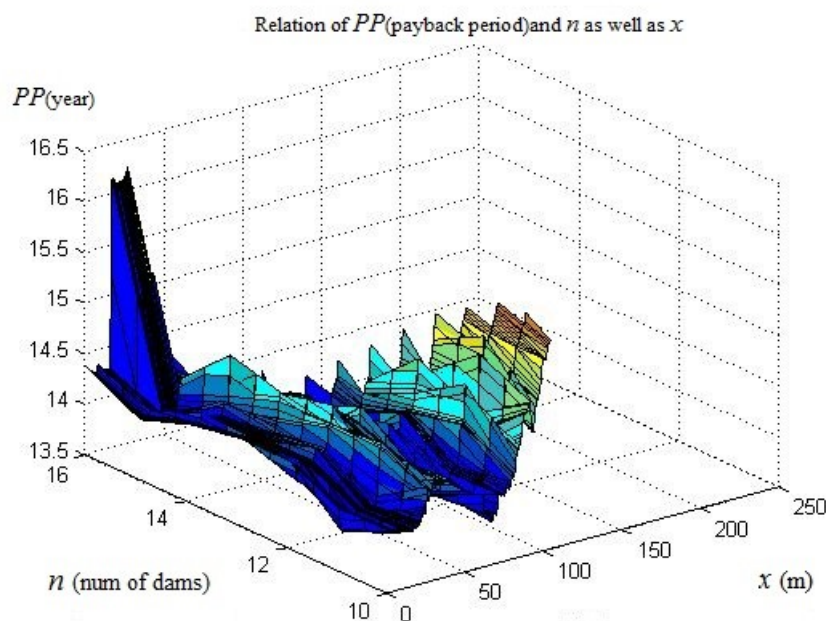

Figure 9: *Relation of $PP(PaybackPeroid)$ and n as well as x*

2. **Calculate all dimensions in the best situation**

   With two unknowns solved, we can figure out all dimensions of small dams. After the calculation, we draw a conclusion that the number of dams is 11, and moreover, the distance between the most upstream small dam and the Zero Point is 45 kilometers. In such a situation, the static payback period is 13.8 years finally. Other dimensions can be calculated as below:

Table 5: three parameters of all small designed dams

$X - axis/km$	45	81	97	113	135	156	187	233	278	346	427
$Dam\ Height/m$	40.87	86.76	125.00	83.38	78.19	67.41	52.39	17.97	33.40	25.33	22.53
$Water\ Level/m$	25.55	54.22	78.13	52.11	48.87	42.13	32.74	11.23	20.87	15.83	14.08

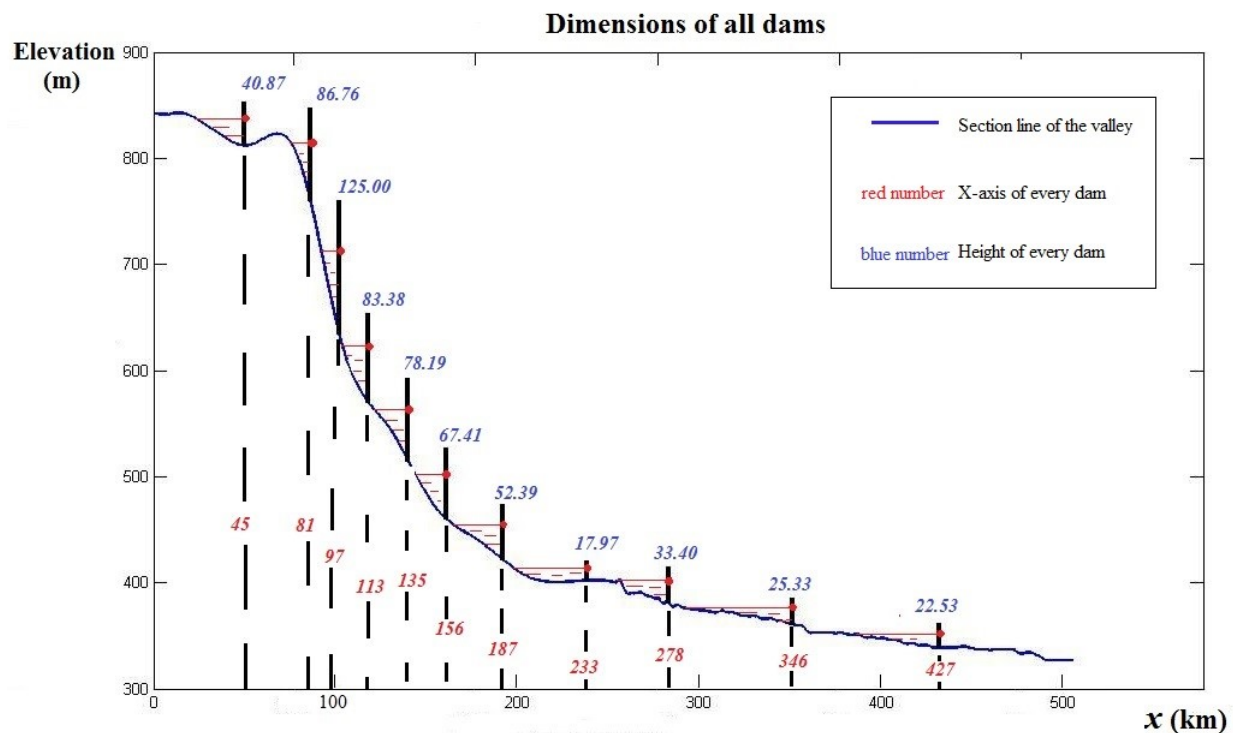Meanwhile, the distribution of all dams can be expressed as following:

Figure 10: Dimensions of all dams

5   Flow Design Model

5.1   Introduction

In order to consider the most detrimental effects of the extreme conditions,we need to predict the maximum discharge and minimum discharge upstream the dams. We choose the probability statistic method to predict that.

Probability statistic method are used to establish a relationship between the random variable and corresponding frequency.

In this model, we take the annually discharge Q as random variable, and calculate the corresponding experienced frequency .Then we try to find a appropriate probability distribution to express the distribution of discharge series.Thus ,we can use this distribution to calculate the discharge value of specified frequency.

5.2   Assumptions

(1) The climate condition is stable, which means that the climate doesnt result in the change of the flow law.

(2) The intensity of Human activities is not large enough to change the underlying surface conditions of the river.

5.3 Model 1: Use Person Type III Distribution to Predict the Maximum Discharge

1. **Choose the discharge series**

   It is obvious in real life that the reservoir capacity determines the scale of the dam, and so dose the benefit. The larger the reservoir capacity usually means more benefit because we will have more water to use.

   We get the annually maximum discharge data (1984-2013) from Global Runoff Data Center(GRDC) as shown in table 6. [4]

Table 6: Annual Maximum Discharge of Zambize River(1984-2013)

year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
discharge(m^3/s)	2275	2337	2898	2609	2831	11500	1184	2957	1037	3927

year	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
discharge(m^3/s)	2319	1762	974	1916	4125	3693	3793	9917	2154	3811

year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
discharge(m^3/s)	5255	3808	3286	10023	9823	5750	9817	5332	3949	4191

We assume that the climate condition is stable ,and the impact of human activities can be negligible. Thus,the flow law of the basin is stable,we can believe that data from 1984 to 2013 are representative enough to reflect the flow law of basin.

2. **Calculate the experienced frequency**

   We take annually maximum discharge Q as random variable, Q obey the distribution function $F(x) = P(Q \geq x)$. $F(x)$means the probability of occurrence of event $Q \geq x$.

   We sort the data by discharge in descending order:

Table 7: Cost Data of Some Dams

sequence number	1	2	3	4	5	6	7	8	9	10
year	1989	2007	2001	2008	2010	2009	2011	2004	2013	1998
max discharge(m^3/s)	11500	10023	9917	9823	9817	5750	5332	5255	4191	4125
frequency(%)	3.23	6.45	9.68	12.90	16.13	19.35	22.58	25.81	29.03	32.26

sequence number	11	12	13	14	15	16	17	18	19	20
year	2012	1993	2003	2005	2000	1999	2006	1991	1986	1988
max discharge(m^3/s)	3949	3927	3811	3808	3793	3693	3286	2957	2898	2832
frequency(%)	35.48	38.71	41.94	45.16	48.39	51.61	54.84	58.06	61.29	64.52

sequence number	21	22	23	24	25	26	27	28	29	30
year	1987	1985	1994	1984	2002	1997	1995	1990	1992	1996
max discharge(m^3/s)	2609	2337	2319	2275	2154	1916	1762	1184	1037	974
frequency(%)	67.74	70.97	74.19	77.42	80.65	83.87	87.10	90.32	93.55	96.77

Then we can calculate the experienced frequency by the following formula:

$$P = \frac{m}{n}$$

where m is the sequence number of discharge data, n is the total number of discharge data. Thus,we find something unreasonable,that the frequency corresponding to the smallest discharge data is 100%. This can not happen absolutely ,so we must make some adjustments. We find that using following formula to calculate the experienced frequency is more reasonable:

$$P = \frac{m}{n+1}$$

We use this formula to work out the experienced frequency corresponding to different discharge,which are listed in the table above.

3. **Choose a mathematical probability distribution model**

   Since we work out the frequency corresponding to different discharge,we can plot the $Q \sim P$ curve.
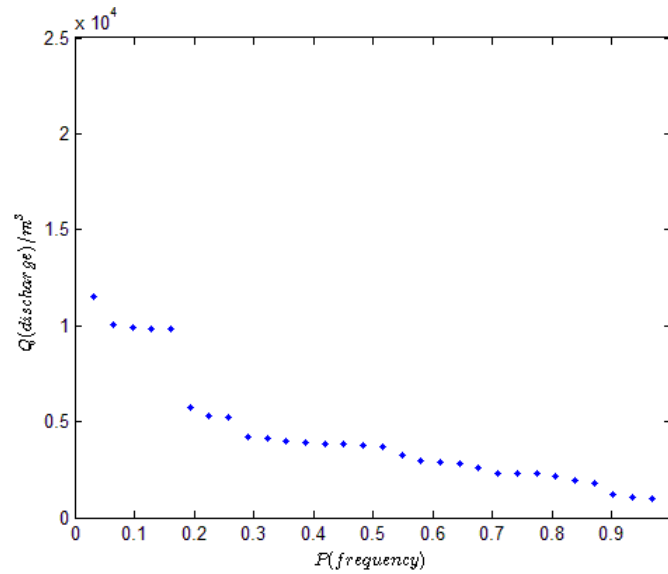

Figure 11: $Q \sim P$ curve

We choose the Pearson type distribution to express the distribution of annually maximum discharge:

$$F(x) = P(Q \geq x) = \int_{x_p}^{\infty} \frac{\beta^\alpha}{\Gamma(\alpha)} (x - \alpha_0)^{\alpha-1} e^{-\beta(x-\alpha_0)} dx$$

where $\Gamma(\alpha)$ is the gamma function of α. α, β, α_0 are three parameters in this distribution. They are related with three statistical parameters : $\bar{x}, C_V, C_S$:

$$\alpha = \frac{4}{C_S^2} \quad \beta = \frac{2}{\bar{x}C_V C_S} \quad \alpha_0 = \bar{x}(1 - \frac{2C_V}{C_S})$$

where $\bar{x}$ is the average of the random variable, C_V is the variation coefficient of the random variable, C_S is the skewness coefficient of the random variable.

We can obtain unbiased estimates of these three statistical parameters by discharge series.

$$C_V = \sqrt{\sum_{i=1}^{n} \frac{(\frac{x_i}{x} - 1)^2}{n-1}} \quad C_S = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{(n-3)\bar{x}^3 C_V^3}$$

By the analysis of data above ,we get $\bar{x} = 4308.5m^3$, $C_V = 0.6848$, $C_S = 1.2877$. Then we can sketch the Pearson-III Curve:
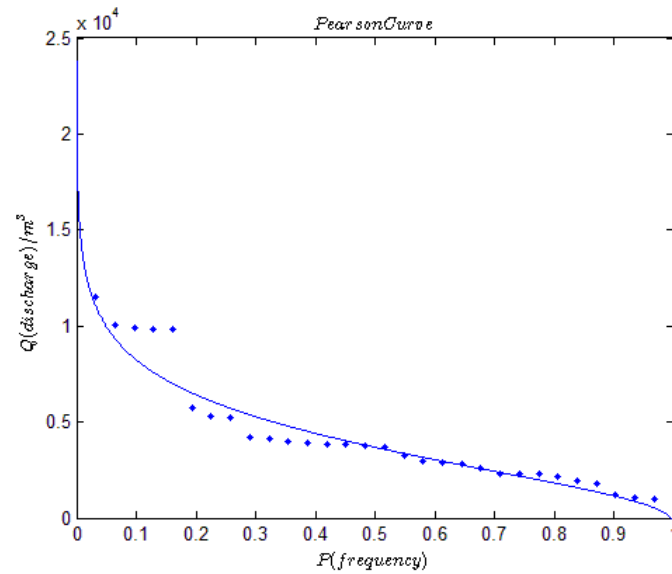

Figure 12: $Pearson - III Curve$

We can see that the discharge variable can be fitted with the Person type distribution.

4. **Predict the maximum discharge by Person type III distribution**

   The *Institute of Risk Management South Africa Risk Research Report* shows that Kariba Dam can resist the magnitude of floodwater that occurs once in 10,000 years. When we replace Kariba dam with smaller dams, we should make sure that they can also resist the flood that occurs once in 10,000 years. Thus, the smaller dams can provide the same level protection. It refers that we need to predict the discharge of the frequency $P = 0.01\%$.Then we have:
   $$0.0001 = \int_{x_p}^{\infty} \frac{\beta^\alpha}{\Gamma(\alpha)}(x - \alpha_0)^{\alpha-1}e^{-\beta(x-\alpha_0)}dx$$

   Solve this equation,then we can get $x_p = 23815m^3/s$,which means that discharge of flood that occurs once in 10,000 years is $23815m^3/s$.We can consider $23815m^3/s$ as the maximum discharge under the extreme condition.

5.4   Model 2: Use Person Type III Distribution to Predict the Minimum Discharge

The prediction process is similar to the above.Here are two differences which we should care about.

1. The formula of experienced frequency turns into $P = 1 - \frac{m}{n+1}$

2. Short-term floodwater can bring tremendous damage.Different from that, short-term low water may mot cause as severe damage,but long-term low water will badly affect people's

livelihood. So we focus on lowest monthly average discharge of each year instead of daily discharge in this model.

Data of the lowest monthly discharge of each year(1984-2013) from GRDC is collected and applied to predict the minimum discharge in this model.Because of the limitation of space, no more tautology here. The Person-III curve is plotted as follows:
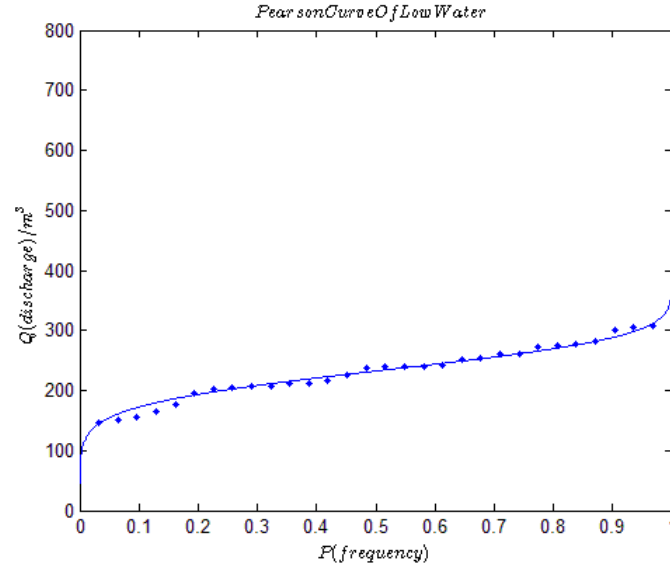

Figure 13: $Pearson - III \; Curve \; Of \; Low \; Water$

Finally, we obtain the monthly discharge of low flow that occurs once in 10,000 year. We can consider the discharge $45m^3/s$ as the monthly minimum discharge under the extreme condition.

6 Connection Reservoirs Operation Model

6.1 Introduction

Through the Decision-Making Model, we obtain x(locations), reservoir capacities and water depths of small dams. These reservoirs are from upstream to downstream on the same river, we call them Connection Reservoirs. Through the Flow Design Model, we obtain the peak flow when flood occurs and the low water flow. Based on these data, we build a model to use them to operate reservoirs.

6.2 Assumptions

1. The operation of a reservoir have impacts on the next reservoirs because they are connected.

2. Operation of connection reservoirs includes pouring water when flood comes and storing water when low water occurs.

3. The Optimal Criterion of operating the connection reservoirs: while running, electric power of reservoirs should achieve maximum value.

4. By changing the order of operating dams, we can get an optimal solution.

6.3   Pouring Model During Flood Period

During flood period, electric power of reservoirs is divided into two parts below.

1. Pouring-power-E_{su}, which is generated while a reservoir pouring water. Set efficiency as η, the average water depth as H, reservoir capacity as the RC, the density of water as ρ, the acceleration of gravity as g, water-storage as V, then:

$$E_{su} = \rho V * g\bar{H} = 9800V\bar{H}$$

   where the unit of E_{su} is W(watt), and $\rho = 1000kg/m^3, g = 9.8N/kg$.

2. Storing-power-E_{st}, which is generated by the inflow of river. Set the inflow of river as W_r, then:

$$E_{st} = \rho W_r g\bar{H} = 9800W_r\bar{H}$$

For connection reservoirs, according to assumption 1, water pouring will hava impacts on E_{st} of the next reservoir. Now consider two reservoirs, as shown in figure 14. Set the id of the upper reservoir as A, another as B.
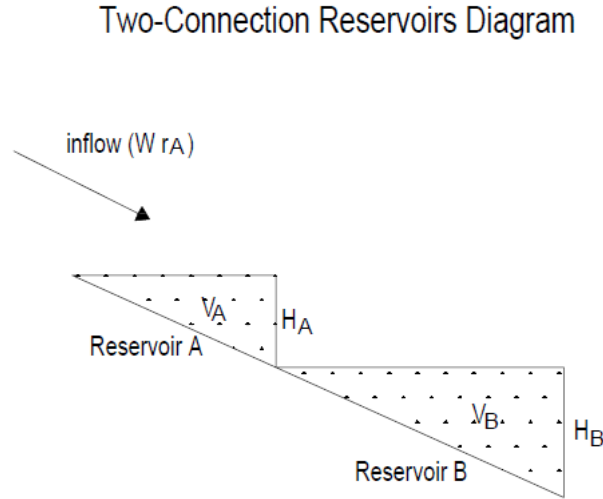

Figure 14: 2-Connection Reservoirs Diagram

Set loss of E_{st} caused by water pouring of A as dE_A, another as dE_B, inflow of river at A as W_{rA}, efficiency of A as η_A, another as η_B, water depth change of A as dH_A, another as dH_B, loss of water-storage of A as V_A, then:

$$dE_A = 9800\eta_A W_{rA} dH_A$$

$$dE_B = 9800\eta_B(W_{rA} + V_A)dH_B$$

If $dE_A < dE_B$, namely $\frac{dH_A}{dH_B} < \frac{W_{rA}+V_A}{W_{rA}}$, then A pouring water before B. Set A pour dV_A, the power of all reservoirs as N, the power of A as N_{VA}, water surface area of A as F_A, then:

$$N_{VA} = 9.8\eta F_A(H_A + H_B)\frac{dH_A}{dt}$$

After water depth of A drops to a secure number, B begins to pour water. Likely, we can get:

$$N_{VB} = 9.8\eta F_B H_B\frac{dH_B}{dt}$$

To ensure the electric power not change(it's determined by users' demond), then:

$$N_{VA} = N_{VB} \Rightarrow F_A(H_A + H_B)\frac{dH_A}{dt} = F_B H_B\frac{dH_B}{dt}$$

Plug it into the former equation, then:

$$\frac{W_{rA}}{F_A(H_A + H_B)} < \frac{W_{rA} + V_A}{F_B H_B}$$

This can be popularized to cases of n reservoirs. Set $k = \frac{W_i}{F_i \sum_{j=i}^{n} H_j}$, where W_i is the inflow of the river plus the water-storage of reservoir i. k can be an index judging the order of pouring water. Reservoir with smaller k should pour water earlier. As our computing result shows, usually upstream reservoir has smaller k, so when flood come, upstream reservoir should pour water earlier.

6.4 Storing Model During Drought Period

Likewise, we can derive another index $K = \frac{W_i'}{F_i \sum_{j=i}^{n} H_j}$, where W_i' is the sum of inflow of the river minus the water-storage of reservoir i. Reservoir with bigger K should store water earlier. As our computing result shows, usually downstream reservoirs have bigger K, so under low water conditions, downstream reservoirs should store water earlier.

6.5 Conclusion

In this model, we assume that the optimal criterion of operating reservoirs is to achieve maximum electric power. Through rigorous mathematical derivation, we get 2 concise indexes for judging the order of pouring and storing water separately. The 2 indexes are easy to use under extreme conditions such as floods and droughts.

7 Strengths and Weaknesses

7.1 Strengths

1. Logical prediction of economic indicators. We devote ourselves to predicting costs and benefits with 7 indexes mentioned above, each of which is linked to the final economic

evaluation directly. Through Back Propagation (BP) neural network, we obtain a precise relationship from 7 indexes or 3 indexes to costs or benefits, and parameter settings are proved to be reasonable from another aspect.

2. Reasonable Decision-Making of dam construction. We focus on the best economic efficiency as the ultimate goal , and take into account two significant unknowns, one is the number of dams, the other is the location of all small dams. In the Decision-Making Model of small dams, we get an optimum solution and provide detailed parameters including the number of dams, coordinates of all dams, all dam heights, water levels of all reservoirs and so on. Furthermore, we adjust theoretical result for a consideration of extreme situation.

3. Adaptable probability statistics model. Person type III distribution is a simple mathematical model which can express the complex hydrological event clearly. Through the curve, we can see that the discharge variable fits well with the distribution. Most importantly, this model can be used to predict discharge of any specific frequency quite accurately.

4. Rigorous mathematical derivation and concise indexes. We use very simple conditions and derive 2 simple indexes for juding the order of pouring water during flood period or storing water during low water period. It's quite useful in operation of connection reservoirs.

7.2   Weaknesses

1. It occurs a range of volatility in the Back Propagation (BP) neural network process, and we need to try many times to get a satisfactory result.

2. Assumptions in the Decision-Making Model may not be totally realistic for increasing assumptions will produce more uncertainties to the final result.

3. In the Flow Design Model, we assume that the flow law is stable over time. But in reality, the climate changes gradually. Global warming does have a impact on the flow law. In addition, human activities may also influence the discharge of the river. Thus, we can consider climate and human activity intensity as parametric variable to obtain more accurate result.

4. The Connection Reservoi Model ignore many factors that will affect the reservoirs in reality, such as the reservoirs may not be connected with local inflow effect. Besides, generate electricity is not the only benefit of reservoirs.

References

[1] Aon.(2015).Impact of the Failure Of The Kariba Dam:The Institude of Risk Management South Africa Risk Research Report

[2] Li, D., Liu, Y., & Chen, Y. (Eds.). (2011). *Computer and Computing Technologies in Agriculture IV: 4th IFIP TC 12 Conference, CCTA 2010, Nanchang, China, October 22-25, 2010, Part 2, Selected Papers* (Vol. 345). Springer.

[3] https://en.wikipedia.org

[4] http://www.bafg.de/GRDC/EN/03_dtprdcts/32_LTMM/longtermmonthly _node.html;jsessionid=8169029F83BED294D3F6906C12F1B03F.live11293

[5] https://www.google.com/earth/

8 Appendix

8.1 Name of Dams

Table 8: Name of dams

Name	ShortForm
ERTAN	ET
XIAOWAN	XW
JINPING	JP
WUDONGDE	WDD
XILUODU	XLD
LONGYANGXIA	LYX
BAIHETAN	BHT
LAWAXI	LWX
LIJIAXIA	LJX