Hydrology and Flow Forecasting

Final Research Report of Cluster 2

The second period of the Yellow River Group Project, a subproject of the China-DC WRE project

June 2002
Hydrology and Flow Forecasting

Researchers: Miss Pang Hui (YRCC)
Ms. Rao Suqiu (YRCC)
Mr. Wang Guoqing (YRCC)
Mr. Huang Xianqing (HU)

Supervisors: Prof. Han Vrijling (DUT)
Prof. Hans Van Duivendijk (DUT)
Dr. Jaap Kwadijk (DUT / WL|Delft Hydraulics)
Dr. Pieter Van Gelder (DUT)

Support: Mr. Wang Wen (PhD student at DUT)

June 5 2002
Delft, the Netherlands
PREFACE

A detailed description of our study and application of the statistical and conceptual models are given in this report. During the three months of staying in the Netherlands, March 12 – June 7, due to the extensive help from the supervisors, the employees of the CICAT, the study and the application can be done successfully. We want express our many thanks for all kindly help for the recommendation of many materials, guidance on our presentation and the final report and so on.
SUMMARY

We have studied and applied the statistic model (i.e. MMC) and hydrological models to Upper Yellow River. This report introduces the results and some conclusions from the model. The three models, MMC, MWBM and NAM, have be applied in the research area.

The forecasted discharge by the three models are closed to the recorded in most low flow months, but in some flood months, the forecasted value is much different with recorded one. And forecasted accuracy by MWBM nearly matches with that by NAM model.

Statistical results indicate that the three models have big forecasted errors in 1989, average relative errors in the year are all excess 60%. In the other four years, MWBM and NAM models have similar forecasted errors; average relative forecasting errors fall in the range from 15% to 55%. For MMC method, average relative forecasted error in 1987 is in smallest value of 23.5% and that in 1990 is in highest value of 102.8%. Forecasted errors in the first two years are relative lower than that in the rest three years, it has increasing trendy.

If acceptable forecasted result is that relative error is less than 20% or absolute error is less than 4mm, then acceptable percentage of forecasted discharge was also calculated. Result shows that most forecasted value by MWBM and NAM is qualified. For MMC method, only in the first two years and low flow period of other years, most forecasted value is acceptable.

Forecasted results by MWBM and NAM are all based on given historical meteorological data. But in real time flow forecasting, forecasting accuracy is mainly up to two factors, one is the hydrological model accuracy and the other is the meteorological forecasting accuracy. And conclusions obtained mainly reflect the first factor.

In addition, in flow forecasting with the statistic method, (i.e. the MMC method), long series are mostly needed. However, 40 years is a reasonable long period of data. The forecasted error in this case is mainly due to not using the rainfall data.

All of results denoted that the three models can be used to forecast discharge in the next one or two years. If meteorological data in the future can be forecasted for a longer period and are of high quality, MWBM and NAM should be adopted first.
Contents

1. INTRODUCTION .................................................................................................................. 6

2. BACKGROUND ..................................................................................................................... 7
   2.1 The Task of Mid-term and Long-term Flow Forecasting .................................................. 7
   2.2 Selection of Study Area ................................................................................................. 9
   2.3 Data Collection ........................................................................................................... 9

3. GENERAL INFORMATION OF STUDY AREA .................................................................. 10
   3.1 Brief Introduction of the YR ....................................................................................... 10
   3.2 General Information about the Reach above Tangnaihai Station .................................. 11
   3.3 Temporal and Spatial Distribution of the Precipitation ................................................. 13
   3.4 Characteristics of Runoff Variation ............................................................................. 17
   3.5 Information of Longyangxia Reservoir ...................................................................... 18
   3.6 Stability Analysis on Runoff Series .......................................................................... 19
   3.7 Frequency Analysis on Discharge .............................................................................. 22
      3.7.1 Introduction ......................................................................................................... 22
      3.7.2 Pearson Type III Distribution ............................................................................ 23
      3.7.3 Fitting the Distribution ....................................................................................... 23
      3.7.4 Test of Pearson Type III Distribution .................................................................. 27

4 REVIEW OF METHODS FOR MID-TERM AND LONG-TERM FLOW FORECASTING .......... 28
   4.1 Definitions .................................................................................................................... 28
   4.2 Forecast Accuracy ........................................................................................................ 28
   4.3 Review of Hydrological Models for Flow Forecasting ................................................... 29
      4.3.1 Index-variable method ......................................................................................... 31
      4.3.2 Storage accounting method ................................................................................ 32
      4.3.3 Extended stream-flow prediction method .............................................................. 32
      4.3.4 Time-series forecasting methods ......................................................................... 32
      4.3.5 Updating long-term forecasts ............................................................................. 32
   4.4 Review on Distributed Physically Based Hydrological Models .................................... 33

5 INTRODUCTION OF THE SOFTWARE IN THE NETHERLANDS ........................................ 35
   5.1 Duflow ......................................................................................................................... 35
      5.1.1 General introduction of Duflow ............................................................................ 35
      5.1.2 RAM model ......................................................................................................... 35
      5.1.2 Application of Duflow ......................................................................................... 39
   5.2 PcRaster ....................................................................................................................... 42
      5.2.1 Introduction of Water Balance Model Based on Mechanism of Saturate Excess Runoff Generation in PcRaster Software Package ..................................................... 42
      5.2.2 Application of PcRaster Software for Discharge Simulation .................................. 43
   5.3 Caterpillar ..................................................................................................................... 45
1. INTRODUCTION

The Yellow River is the second longest river in China, and it is the vital water source of north and northwest of China. Originating from the northern foot of the Bayankela Mountain in Qinghai Province, it runs through 9 provinces and autonomous regions and flows into the Bohai Sea in Shandong Province. It is 5,464 km long with a basin of 752,000 km$^2$. The mean runoff volume in years is about 58 billion m$^3$. Because of development of the industry and agriculture along the river, the water consumption increases rapidly. The shortage of water resources becomes very serious in Yellow River, especially in middle and lower reach. It caused the zero flow occurred frequently in the lower reach in 1990s. There are several successive dry years in this period. During the period, the zero-flow occurs earlier and earlier, the length of the dry reach and the duration of zero-flow became longer and longer. The length of the dry reach is even more than 700 km, about one eighth of the total river. In 1997, Lijin station, the last station of the main stream, has had zero-flow up to 226 days. At that time in the estuary it's near 300 days that no water flows into the sea. Water resources scarcity and dry river not only bring trouble to industry and agriculture, but also further deteriorate the environment and the downstream channel. From 2000, because of more efficiently integrated water resources management, the zero-flow situation doesn't occur any longer.

Chinese government decided to carry out the integrated water resources management in yellow River in 1998, and YRCC is with responsibility for this work. For the operation of water resources YRCC demands mid-term and long-term flow forecasts of the discharge of main stations in the trunk and some branches of YR basin urgently. Moreover, many reservoirs in the upper and middle reach for hydropower generation, irrigation and prevention from ice flood of frozen reach in winter have to use the mid-term and long-term flow forecasts also.

The mid-term and long-term flow forecasting models are urgently needed in YRCC. In past two years we had many works to settle the data and developed some relative models, but it can’t satisfy the demand far off. One of the targets of China-Dutch WRE Project is to study hydrology and flow forecasting models and apply the technology in the Yellow River basin.
2. BACKGROUND

2.1 The Task of Mid-term and Long-term Flow Forecasting

According to demand, the forecast department of YRCC must supply the different forecasting productions at different date in time. The task table is as follow. The station distributes as figure2-1.

Table 2-1. The mid- and long-term flow forecasting task of Yellow River in low flow season

<table>
<thead>
<tr>
<th>Bulletin Date</th>
<th>Item</th>
<th>Station</th>
<th>Volume from Dec. to Jun.</th>
<th>Monthly discharge from Dec. to Jun.</th>
<th>Monthly discharge of next month</th>
<th>Mean discharge of next period of ten days</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 October</td>
<td>25 of each month</td>
<td>The eighth day of each period of ten days</td>
<td>Tangnaihai</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Zheqiao</td>
<td></td>
<td></td>
<td>HKongqi</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Minghe</td>
<td></td>
<td></td>
<td>HKiangtang</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Longmen</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Hejin</td>
<td></td>
<td></td>
<td>Zhuangtou</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Huaxian</td>
<td></td>
<td></td>
<td>Tongguan</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Heishiguan</td>
<td></td>
<td></td>
<td>Wuzhi</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>
Fig. 2-1. The Yellow River Basin
It's very difficult to forecast the flow of Yellow River basin because there are many factors which affect the variation of flow. Like precipitation, temperature, river freezing, operation of reservoirs, irrigation, water and soil conservation in middle area, and so on. Some models are developed for the upper and middle stream, which is based on the statistics correlative analysis. But further studies must be made for adapting the demands of practice.

2.2 Selection of Study Area

In this research Tangnaihai station, in the upper stream of YRB, is selected as the study area. The main reasons are as follow.

1. In the dry season, the most water comes from the headwaters of Lanzhou station of the upper reach, and water volume upstream Tangnaihai occupies 56% of total at Huayuankou station which is the last one of the upper and middle stream in the Yellow River basin.

2. Tangnaihai station is the inflow station of Longyangxia reservoir, which is the large reservoir with about 25 billion m$^3$ of storage capacity at normal storage water level at the most upper of YR. The runoff forecasting is very important to the operation and generation of the reservoir.

3. Upwards Tangnaihai there is sparsely populated and it is affected by little human activities. The flow productive can be considered as “natural”. Then it is more convenient to use some models for flow simulation in this area.

2.3 Data Collection

The data was collected from the research area for developing the models. It includes precipitation, evaporation, temperature, discharge, section profile and the information of Longyangxia reservoir. The data information lists in table 2-2.

**Table 2-2. Data list**

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>Number of stations</th>
<th>Data types</th>
<th>filename</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Precipitation</td>
<td>13</td>
<td>Monthly and ten days</td>
<td>*.dbf, *.xls</td>
</tr>
<tr>
<td>2</td>
<td>Runoff</td>
<td>4</td>
<td>Monthly</td>
<td>*.xls</td>
</tr>
<tr>
<td>3</td>
<td>Data of Section</td>
<td>1</td>
<td>Data observed in 1997</td>
<td>*.xls</td>
</tr>
<tr>
<td>4</td>
<td>Graph</td>
<td>1</td>
<td>Map of Upper reach above Tangnaihai</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Data of Reservoir</td>
<td>1</td>
<td>Storage, water level, Curves, and power ability, etc</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Evaporation</td>
<td>1</td>
<td>Monthly</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Temperature</td>
<td>1</td>
<td>Monthly</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>General information</td>
<td>The basin upward Tangnaihai</td>
<td>Latitude, longitude and the distance between stations, etc</td>
<td></td>
</tr>
</tbody>
</table>
3. GENERAL INFORMATION OF STUDY AREA

3.1 Brief Introduction of the YR

Yellow River is the second longest river in China with the length of 5464 km. Compared with other rivers in China, YR is characterized with such following points: poor in water resources and rich in sediment, uneven distribution in time and space, suspended river, frequent flood and severe damage.

The precipitation situation of Yellow River: Most of precipitation in YR basin is in form of rain. The precipitation in snow is a very small part of the whole. The multi-year average precipitation is 370.1 billion m³, only 6 percent of that of the whole country. The multi-year average annual precipitation in depth is 465mm. Generally the precipitation is decreasing gradually from Southeast to Northwest. The largest precipitation occurs in the north slope of the Qinling Mountain. The multi-year average precipitation can be up to over 900mm in some area, and the least precipitation occurs in Hetao area of Ningxia and Inner Mongolia with only 200-300mm. Timely distribution is uneven with a year. Most of precipitation occurs in the summer (June-August), i.e more than 50% of annual precipitation.

The temperature situation of YR: The temperature in Southeast is higher than that of the Northwest. The highest average monthly temperature is often in July, in most area, it’s between 20°C~29°C. The highest temperature is even high to over 40 °C in some area. The lowest temperature is often in January, in most areas it’s below 0 °C.

The Water resources of YR: Although it is the second longest river in China, it’s not rich in water resources. The multi-year average natural runoff is 58 billion m³, only about 2.1% of that of the whole of China. The average runoff depth is about 70 mm, less than 30% of that of the whole of China. Average water quantity per person is only 593m³, about 23% of the whole country.

The spatial distribution of the runoff is uneven in YR basin. The reach upper from Lanzhou is the most abundant water resources.

Generally YR is divided into three parts. The upper reach is from the river source to Toudaoguai. The middle reach is from Toudaoguai to Huayuankou and the lower reach is from Huayuankou to the estuary.

According to basin characters, river system and the situation of flow concentration, the upper reach can be divided into four subbasin. From upstream to downstream, they are the reach upward Tangnaihai, from Longyangxia to Liujiaxia, from Liujiaxia to Lanzhou, from and Lanzhou to Toudaoguai.
3.2 General Information about the Reach above Tangnaihai Station

The area of the reach up from Tangnaihai is 121917 km$^2$. The length of main stem is over 1500 km, near one fourth of the total river length. Multi-year average runoff is about 21 billion m$^3$, and average value of runoff during low-flow season is 8.3 billion m$^3$, about 40% of the annual value.

13 rain gages and 3 hydrologic stations are located in the area. In addition, there is one meteorology station being used to collect the evaporation and temperature data.

The basic information about the stations is shown in the following tables. (from upstream to downstream)

Table 3-2 shows the information about the hydrology stations. Table 3-1 and 3-3 show the information about the rainfall stations. Figure 3-1 shows the sketch map of research area and the distribution of rainfall gages and hydrology stations in this area.

Table 3-1. Rainfall Station name and duration of data

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Name</th>
<th>Duration of data</th>
<th>Nr.</th>
<th>Name</th>
<th>Duration of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Marqu</td>
<td>1964-1998</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3-2. Information of the Hydrologic stations

<table>
<thead>
<tr>
<th>Station</th>
<th>Longitude (East)</th>
<th>Latitude (North)</th>
<th>Distance to river mouth (km)</th>
<th>Drain area (km$^2$)</th>
<th>Duration of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marqu</td>
<td>102°05’</td>
<td>33°58’</td>
<td>4284</td>
<td>86048</td>
<td>1960-1997</td>
</tr>
<tr>
<td>Jungong</td>
<td>100°39’</td>
<td>34°42’</td>
<td>4057</td>
<td>98414</td>
<td>1980-1997</td>
</tr>
<tr>
<td>Tangnaihe</td>
<td>100°09’</td>
<td>35°30’</td>
<td>3911</td>
<td>121972</td>
<td>1956-2001</td>
</tr>
</tbody>
</table>
Table 3-3. Information of the Rainfall stations

<table>
<thead>
<tr>
<th>Station</th>
<th>Longitude (East)</th>
<th>Latitude (North)</th>
<th>Station</th>
<th>Longitude (East)</th>
<th>Latitude (North)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jimai</td>
<td>99°39'</td>
<td>33°46'</td>
<td>Jiuzhi</td>
<td>101°30'</td>
<td>33°26'</td>
</tr>
<tr>
<td>Mentang</td>
<td>101°03'</td>
<td>33°46'</td>
<td>Xinghai</td>
<td>99°58'</td>
<td>35°36'</td>
</tr>
<tr>
<td>Tangke</td>
<td>102°28'</td>
<td>33°25'</td>
<td>Henan</td>
<td>101°36'</td>
<td>34°42'</td>
</tr>
<tr>
<td>RuoerGai</td>
<td>102°56'</td>
<td>33°35'</td>
<td>Zeku</td>
<td>101°29'</td>
<td>35°03'</td>
</tr>
<tr>
<td>Maqiu</td>
<td>102°05'</td>
<td>33°58'</td>
<td>Guoluo</td>
<td>100°16'</td>
<td>34°30'</td>
</tr>
<tr>
<td>Jungong</td>
<td>100°39'</td>
<td>34°42'</td>
<td>Renxiangmu</td>
<td>99°13'</td>
<td>34°08'</td>
</tr>
<tr>
<td>Tangnihe</td>
<td>100°09'</td>
<td>35°30'</td>
<td>Hongyuan</td>
<td>102°10'</td>
<td>32°48'</td>
</tr>
</tbody>
</table>

Fig. 3-1. Sketch map of research area upward Tangnaihai

Tangnaihai station is located at the upper stream of Longyangxia reservoir, the river at this reach runs through a great gorge. The section of Tangnaihai station is very regular because the river bed and two sides banks consist almost completely of stone. Figure 3-2 shows the profile of the section measured on April 25, 1997.
3.3 Temporal and Spatial Distribution of the Precipitation

The locations of 13 rainfall stations are shown in figure 3-1. Multi-year average precipitation of each station is shown in figure 3-3. Comprehensive analysis indicates that those stations located in Southeast have more precipitation than those in Northwest. It is also uneven in space distribution with a tendency of decrement from Southeast to Northwest. The maximum multi-year average precipitation is 778mm in Jiuzhi station while the minimum one is only 316mm in Xinghai station.

Annual area precipitation in the study area and its trend line from 1964 to 1990 are shown in fig 3-4. Annual precipitation curves of each station are shown in the fig3-5 to fig3-8. Fig 3-4 shows that annual average area precipitation has decreasing tendency. And annual precipitation is uneven from year to year. Those successive poor or rich years last only two or three years in general. Maximum annual precipitation took place in 1967 with depth of 715.1mm and the minimum annual precipitation occurs in 1990 with depth of 467.2mm.
Fig 3-3. Multi-year average precipitation at each station

Fig 3-4. The area average precipitation curve
Fig 3-5. Annual precipitation of Jimai, Jiuzhi, Renxiamu and Mentang

Fig 3-6. Annual precipitation curve of Henan, Zeku and Xinghai
Fig 3-7. Annual precipitation of Tangke, Hongyuan, Ruoergai and Marqu

Fig 3-8 Annual precipitation curve of Guoluo and Jungong
3.4 Characteristics of Runoff Variation

Figure 3-9 shows the yearly water volume from 1956 to 2001 at Tangnaihai station. The average value of multi-year is about 20 billion m$^3$. The maximum annual runoff occurs in 1989 with 32.8 billion m$^3$ and the minimum one occurs in 1956 with 13.3 billion m$^3$. Analysis indicates the research area has been in the recent relatively drought period which lasted more than 10 years. Only two years the annual runoff is more than the multi-year average value since 1990. They are in 1993 with 21.9 billion m$^3$ water volume and 1999 with 24.3 billion m$^3$ and only a little more than the average value.

![Fig. 3-9. The water volume of year at Tangnaihai station](image)

![Fig. 3-10. Monthly average discharge at Tangnaihai station](image)
The discharge at Tangnaihai station has a large variation for each month. Figure 3-10 shows the monthly average discharge curve at Tangnaihai station. The maximum discharge appears in July, the second one appears in September. The minimum one appears in January. The runoff during low flow season from November to June is 40% of total of the year.

Figure 3-11 shows the yearly average runoff variation at each station upward Tangnaihai. The multi-year average runoff of Jungong, Maqu and Jimai are separately 89%, 83% and 28% of that of Tangnaihai.

![Fig. 3-11. Yearly average runoff variation upward Tangnaihai](image)

### 3.5 Information of Longyangxia Reservoir

Longyangxia reservoir is the largest reservoir except Xiaolangdi reservoir being built in Yellow River basin. Its designed maximum storage is over 30 billion m$^3$. The technical economic indexes of Longyangxia power station are as following:

- Normal storage water level: 2600m.
- Storage capacity at normal storage water level: 24.7km$^3$.
- Regulation storage: 19.35km$^3$.
- Regulation property: Pluriennial.
- Installed capacity: 1280MW.
- Firm output: 589.8MW
- Annual energy output: 5.942 *10$^8$ TW.h
The feature curves are shown as figure 3-12 and 3-13.

Fig. 3-12. Water level - storage curve of Longyangxia reservoir

Fig. 3-13. Water level - area curve of Longyangxia reservoir

3.6 Stability Analysis on Runoff Series

Stability analysis on runoff series is conducted to detect whether basin hydrology is impacted by human activities. If basin is disturbed by many non-climate factors obviously, such as large-scale water and soil conservation measures are taken in a basin or hydraulic projects are build up in main stem. Then, corresponding responses would occur on the runoff series.
Because month-to-month values vary more widely than year-to-year values even in the natural situation, annual data can be used to analyse stability of runoff series. To eliminate the sharp fluctuations in the annual value, years moving average of time series is often adopted to carry out the task.

The reach above Tangnaihai station is chosen as the study area. The area is located in the upper reaches of Yellow River. Yearly discharge and 5-year moving average process at the gauged station are shown in Fig.3-14 and annual runoff characteristics values in different period are given in Table.3-4.

![Fig.3-14 Yearly discharge and 5-year average process at Tangnaihai station from 1956 to 1989](image)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>discharge</td>
<td>164.6</td>
<td>177.5</td>
<td>165.0</td>
<td>203.9</td>
<td>177.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year moving average</td>
<td>265.3</td>
<td>254.9</td>
<td>254.1</td>
<td>257.9</td>
<td>265.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table.3-4. Annual runoff depth (mm) characteristics value at Tangnaihai station in different periods**

If a time series were stable, then all the points should vary in a limited range and distribute around the multi-year value, No obvious jump appeared. Fig3-16 shows that all points distribute around the multi-year average discharge line and have slightly increasing trend. Results in Table3-4 also show that average value of discharge from 1956 to 1986 is
177.8 mm, nearly equal to that in the period from 1960 to 1969. And multi-year average value is slightly larger than mean values in 1950' and 1970', but slightly less than that in 1980'. The maximum value in each period varies in a limit range from 254.1 to 265.3, and the minimum value varies only from 117.3 to 155.3.

In addition, double accumulative curve method is usually applied to analyse stability of time series. If no points diverge much from main trend line, then one can conclude the series is stable. Double accumulative curve for annual discharge at Tangnaihai station versus to average area precipitation was plotted in fig3-15. And the figure shows that all points are all around in a perfect tend line.

![Double accumulative curve](image)

**Fig.3-15 Double accumulative curve for annual discharge versus to annual precipitation**

Difference between Relationships of precipitation and discharge in different periods also can reflect series' stability to some extent. Series from 1956 to 1986 was separated into two parts. One is from 1956 to 1970; the other is from 1971 to 1986. Linear relationship between annual precipitation and discharge in the two periods was established based on the correlation analysis. And best-fit line is shown in Fig.3-16. The similar best-fit lines in different period indicated there is no obvious difference between the relationships in difference periods. And the distribution of the points in the figure shows that all points in different periods concentrate in a domain.
Fig. 3.16 Relationship between annual precipitation and discharge in different period

The hydrological time series, as a sample of a data population, must be first considered in relation to the statistical population of that variable. If the statistics of the sample (mean, variance, etc) are not functions of the timing or length of the sample, then the time series is said to be stationary. If a definite trend is discernible in the series, then it is non-stationary series. Similarly, periodicity in a series means that it is non-stationary. The modelling of a time series is much easier if it is stationary, so identification, quantification and removal of any non-stationary component in a data series is undertaken, leaving a stationary series to be modelled.

3.7 Frequency Analysis on Discharge

Many of the statistical techniques that are applied to hydrologic data (to enable inferences to be made about particular attributes of the data) can be labeled with the term “frequency analysis” techniques. The term “frequency” usually connotes a count (number) of events of a certain magnitude. To have a perspective of the importance of the count, the total number of events (sample size) must also be known. The probability of a certain magnitude event recurring again in the future, if the variable describing the events is continuous, (as are most hydrologic variables), is near zero. Therefore, it is necessary to establish class intervals (arbitrary subdivisions of the range) and define the frequency as the number of events that occur within a class interval.

3.7.1 Introduction

Benson (1968) reported on a method of flood frequency analysis based on the Pearson type III distribution that is obtained when the logs of observed data are used along with the Pearson type III distribution. Pearson type III (P-III) distribution was recommended by Chinese national design flood norm and curve-fitting method by eye estimation was
recommended by Chinese national design flood norm also as a parameter estimation method, so, this method is applied for the monthly discharge data of Tannaihai basin.

### 3.7.2 Pearson Type III Distribution

General. There are 4 hydrological Stations in Tannaihai basin, and the discharge data of 4 Hydrological Stations are shown as fellows.

1. Tonlaihe Station (1956-2000)
2. Jimai Station (1959-1997)
3. Maqu Station (1960-1997)

The distribution of average monthly discharge of 4 Stations is shown in Fig.3-17.

![Graph showing the distribution of average monthly discharge of 4 stations](image)

**Fig. 3-17 The distribution of average monthly discharge of 4 stations**

### 3.7.3 Fitting the Distribution

(1) The Pearson type III distribution is fitted to a data set by calculating the mean monthly discharge, the coefficient of variance, and the coefficient of skewness from the following equations:

\[
EX = \frac{\sum X}{N}
\]

\[
\sigma = \sqrt{\frac{E(X - \bar{x})^2}{N}}
\]

\[
Cv = \frac{\sigma}{EX}
\]
\[ Cs = \frac{E(X - \bar{x})^3}{\sigma^3} \]

In which:

- \( EX \) = the magnitude of the monthly discharge
- \( N \) = number of months in the data set
- \( \sigma \) = the standard deviation of the variance
- \( Cv \) = the coefficient of variance
- \( Cs \) = the coefficient of skewness

The precision of the computed values is more sensitive to the number of significant digits.

(2) In terms of the frequency curve itself, the mean represents the general magnitude or average ordinate of the curve, the square root of the variance (the standard deviation, \( \sigma \)) represents the slope of the curve, and the skew represents the degree of curvature. Computation of the unadjusted frequency curve is accomplished by computing values for the monthly discharge corresponding to selected values of percent chance exceedance.

Pearson type III distribution with 4 hydrological stations are shown in following:

Fig.3-18. PIII Distribution with Tangnaihai
Fig. 3-19. PIII Distribution with Jimai

Fig. 3-20 PIII Distribution with Maqu
A reasonable set of values and the results are shown in Table 3-5.

Table 3-5. Results of Curve-fitting method

<table>
<thead>
<tr>
<th>Station</th>
<th>Ex</th>
<th>Cv</th>
<th>Cs</th>
<th>$Q(0.001%)$ (m$^3$/s)</th>
<th>$Q(0.01%)$ (m$^3$/s)</th>
<th>$Q(0.1%)$ (m$^3$/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jimai</td>
<td>129.5</td>
<td>0.9</td>
<td>2.25</td>
<td>1445.3</td>
<td>1150.5</td>
<td>855.6</td>
</tr>
<tr>
<td>Maqu</td>
<td>469.5</td>
<td>0.78</td>
<td>1.95</td>
<td>4259.8</td>
<td>3435.8</td>
<td>2608.2</td>
</tr>
<tr>
<td>Jungong</td>
<td>568</td>
<td>0.8</td>
<td>2</td>
<td>5343.7</td>
<td>4298.6</td>
<td>3253.5</td>
</tr>
<tr>
<td>Tangnaihai</td>
<td>638.8</td>
<td>0.76</td>
<td>1.9</td>
<td>5590.8</td>
<td>4513</td>
<td>3440.1</td>
</tr>
</tbody>
</table>

Some conclusions

(1) Monthly flow processes at Tangnaihai is similar to the other three hydrological stations, it shows strong seasonality

(2) The same trend: two peaks in each year in average monthly flow processes of all four stations.

(3) Obtain the maximum monthly discharge of Yellow River basin at four stations with 1000, 100, 50 year return period

Discussion

There is a problem in fitting the curves, which is why the value of $Cv$ of monthly discharge in Jimai station is larger than the other three stations. (see table 3-5.)

The reasons include two aspects:

(1) Curve-fitting method by eye estimation has large randomness.
Curve-fitting method by eye estimation has some advantages, such as large flexibility, may consider different expert opinions, easy to do adjust in time and space. But it has some disadvantages, for example, large randomicity; different person may have different estimation result. The result may be changed for different purposes.

(2). No extraordinary larger monthly discharge value only in Jimai station. It has impacts to the fitting of curve, and the inducing the change of variables.

3.7.4 Test of Pearson Type III Distribution

The soundness of a probability distribution can be tested by comparing the theoretical and sample values of relative frequency or the cumulative frequency function. In the case of the relative frequency function, the \( X^2 \) test is used. The test of Pearson type III distribution by Pearson \( X^2 \) Method (just use in Tannaihai station) is as follows:

\[
E_x = 638.8 \text{ (m}^3/\text{s)}
\]
\[
C_v = 0.76
\]
\[
C_s = 1.9
\]

Adopt \( K = 6 \),
\[
\pi_i = 1/6 = 0.1667, \ n = 549,
\]

So. \( n\pi_i = 91.5, \ i = 1-6 \)

The compute result of \( X^2 \) test is shown in Table 3-6.

<table>
<thead>
<tr>
<th>Table 3-6. Compute result of ( X^2 ) test</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>( f_i )</td>
</tr>
<tr>
<td>( n\pi_i )</td>
</tr>
<tr>
<td>( f_i - n\pi_i )</td>
</tr>
</tbody>
</table>

Hypothesis:

A confidence level is chosen for the test: \( 1 - \alpha = 0.05 \)

\[
X^2 = \sum \frac{(f_i - n\pi_i)^2}{n\pi_i} = \frac{400.4725}{91.5} = 4.377
\]

and

\[
X^2 < X^2_{1-\alpha} = 5.991
\]

So this hypothesis is accepted.
4 REVIEW OF METHODS FOR MID-TERM AND LONG-TERM FLOW FORECASTING

4.1 Definitions

Forecasting means the estimation at a specific time, or during a specific time interval. It's distinguished from prediction, which is the estimation of future conditions, without reference to a specific time. Long term forecast are those with long lead times, usually up to several months, at present, little forecast skill is possible for hydrologic variables when forecast lead time extend beyond several months to a year; for longer lead times, prediction techniques are more applicable. Seven days is a convenient division between short- and long-term forecasts because there is slightly longer than the current maximum lead times of 96h for the U.S. national weather service's quantitative precipitation forecast.

4.2 Forecast Accuracy

Forecast accuracy is a measure of difference between the amount forecasted and the value that actually occurs. Forecast errors can be either systematic or random. For forecast period T in year I, \( I = 1, 2 \ldots \) the means of the forecasts and observation can be defined as follow:

\[
M_f = \frac{1}{n} \sum_{i=1}^{n} Q_f(i) \quad M_o = \frac{1}{n} \sum_{i=1}^{n} Q_o(i)
\]

Where \( M_f \) and \( M_o \) are mean value of forecasted and observed discharge respectively, \( Q_f(i) \) and \( Q_o(i) \) are forecasted and observed discharge at the forecast time \( i \) respectively.

The following are widely used measures of forecast errors:

Bias: \( B = M_f - M_o \)

Mean squared error: \( MSE = \frac{1}{n} \sum_{i=1}^{n} [Q_f(i) - Q_o(i)]^2 \)

Root mean square errors: \( RMSE = (MSE)^{0.5} \)

Variance: \( V = MSE - B^2 \)

Relative bias: \( RB = \frac{B}{M_o} \)

Mean absolute error: \( MAE = \frac{1}{n} \sum_{i=1}^{n} |Q_f(i) - Q_o(i)| \)

Relative mean absolute error: \( RMAE = \frac{MAE}{M_o} \)
Forecast efficiency: \[ E = 1 - \frac{MSE}{V} \]

R squared: \[ R^2 = \left[ \frac{\frac{1}{n} \sum_{i=1}^{n} Q_f(i)Q_o(i) - M_oM_f}{\left(\frac{1}{n} \sum_{i=1}^{n} Q_o^2(i) - M_o^2\right)^{\frac{1}{2}}} \right] \left[ \frac{1}{n} \sum_{i=1}^{n} Q_f^2(i) - M_f^2 \right] \]

Among the above quantitative standards, bias and relative bias are measures of systematic error in the forecast. Over the period of many years, they measure the degree to which the forecast is consistently above or below the actual value; variance is a measure of the variability, a measure of the random error. Mean square error, root mean square error, mean absolute error, relative mean absolute error and forecast efficiency are all measures that incorporate both systematic and random error. Errors in long-term forecasts arise from three sources: model error, data error and meteorological error.

4.3 Review of Hydrological Models for Flow Forecasting

Hydrological models can be divided broadly into two groups. One is deterministic model, and the other is stochastic model.
<table>
<thead>
<tr>
<th>Country</th>
<th>Model name</th>
<th>Authority</th>
<th>Main purposes</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’DONNELL</td>
<td>Flow forecasting</td>
<td>Dawdy and O’Donnell</td>
<td></td>
</tr>
<tr>
<td>STANFORD</td>
<td>Flow forecasting with snow melt</td>
<td>Linsley and Crawford Stanford University</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>DISPRIN</td>
<td>Dee and Clwyd river Authority</td>
<td>River regulation</td>
</tr>
<tr>
<td>UK</td>
<td>HYRROM</td>
<td>Institute of hydrology</td>
<td>Assess effect of change in land use</td>
</tr>
<tr>
<td>UK</td>
<td>LAMBERT</td>
<td>Dee and Clwyd river Authority</td>
<td>River regulation</td>
</tr>
<tr>
<td>UK</td>
<td>HYSIM</td>
<td>Manley Severn Trent Water Authority</td>
<td>Works design and operation</td>
</tr>
<tr>
<td>Australia</td>
<td>BOUGHTON</td>
<td>Murray</td>
<td>Daily discharge forecasting</td>
</tr>
<tr>
<td>Australia</td>
<td>CBM</td>
<td>Commonwealth Bureau of Meteorology</td>
<td>Forecast flood flows</td>
</tr>
<tr>
<td>Canada</td>
<td>UBC</td>
<td>University of British Columbia</td>
<td>Flow forecasting with snowmelt</td>
</tr>
<tr>
<td>France</td>
<td>BILIK</td>
<td>Sogreah, Grenoble</td>
<td>General forecasting</td>
</tr>
<tr>
<td>France</td>
<td>CREC</td>
<td>Chatou</td>
<td>Discharge forecasting</td>
</tr>
<tr>
<td>France</td>
<td>GIRARD1</td>
<td>ORSTOM, Paris</td>
<td>Multipurpose</td>
</tr>
<tr>
<td>Israel</td>
<td>MERO</td>
<td>TAHAL Eng.Co.</td>
<td>Cyprus water planning</td>
</tr>
<tr>
<td>Italy</td>
<td>CLS</td>
<td>Pavia Univ. and IBM Pisa</td>
<td>Flood forecasting</td>
</tr>
<tr>
<td>Japan</td>
<td>TANK2</td>
<td>Yodo R.Dams Control Hirakate city</td>
<td>Discharge for water resources projects</td>
</tr>
<tr>
<td>South Africa</td>
<td>HYREUN</td>
<td>Hydrological Research Unit, Univ. of Witwatersrand</td>
<td>Flood hydrograph simulation</td>
</tr>
<tr>
<td>USA</td>
<td>NWSH</td>
<td>Nat. Weather Service Maryland</td>
<td>Flood and low flow forecasting</td>
</tr>
<tr>
<td>USA</td>
<td>SRFCH</td>
<td>Nat. Weather Service Sareramento R. forecast center</td>
<td>Flood and low flow forecasting</td>
</tr>
<tr>
<td>USA</td>
<td>SSARR</td>
<td>US Army Corps of Engineers Portland Oregon</td>
<td>Flood and low flow forecasting</td>
</tr>
<tr>
<td>USA</td>
<td>HEC</td>
<td>US Army Corps of Engineers Davis, California</td>
<td>Flood and low flow forecasting</td>
</tr>
<tr>
<td>USSR</td>
<td>HMC</td>
<td>Hydromet. Centre, Moscow</td>
<td>Short-term forecast of floods</td>
</tr>
<tr>
<td>China</td>
<td>XIN’ANJIANG</td>
<td>Zhao Renjun, Hohai University</td>
<td>Flow forecasting</td>
</tr>
</tbody>
</table>

There is clear distinction between the two models. In the wide sense, stochastic model is used to refer to the models of system or subsystems that contain a large random element, and
generally sample randomly from a range of probable values. It describes the hydrological time series of the several measured variables such as rainfall, evaporation and stream flow involving distributions in probability. And the deterministic model is one in which a given input of rain must produce a fixed output of runoff in a certain physical environment, its output is also determined. It seeks to simulate the processes in the catchment involved in the transformation of rainfall to stream flow. Deterministic models can be a very simple empirical model or black box model that derived from input-output analysis of past data. And it also can be conceptual model in which processes are represented in a very simplified manner, perhaps including some empirically calculated elements. Or it can be a physical based model, in which processes are modelled according to fundamental scientific laws. Classification of hydrological model was given in Fig.4-1.

The advent of the high-speed digital computer with a large storage for data has greatly improved the development of conceptual hydrological models. At present, more conceptual models devote themselves to describe several physical component process of hydrological cycle. For simplification, the hydrology of a drainage basin, from precipitation through to a stream discharge at the lowest outlet, can be conceived as a series of inter-linked processes and storage, the storage are considered as reservoirs, for which water budgets are kept. And now, conceptual hydrological models are widely used in planning of water resources, flow forecasting and anther aspects. Some of the most popular models used in the worldwide were listed in the Table 4-1.

Table 4-1 shows that hydrological models were widely used in flow forecasting. At present, long-term forecasting models generally fall into three classes: index-variable, storage accounting and conceptual simulation models. The general used forecasting models as follow:

### 4.3.1 Index-variable method

An index variable is a surrogate, readily measured prior to the time of forecast, which can be related to forecast period runoff. Index-variable methods are of the general form

\[ Q_f = f(X_1, X_2, \ldots, X_n) \]

Where \( X_i \) is the \( i^{th} \) of \( n \) index variable. \( X_i \) might indicate values of the same variable at different measure locations.
4.3.2 Storage accounting method

Storage accounting schemes is that runoff in a future forecast period is determined by the amount of water presently in storage in a catchment, as well as forecast period precipitation. The model can be viewed as a special case of the index-variable method with regression estimate of parameters.

Basin storage could be estimated as a linear function of basin precipitation from the beginning of an accounting period up to the forecast date, less runoff:

\[ S_t = a + b \cdot P_t - R_t \]

Where \( P_t \) is the total precipitation at one or more precipitation gauges and \( R_t \) is the total runoff in the same period. The forecast runoff for the period \( T \) is then a linear function of \( S_t \),

\[ Q_f(t,T) = \alpha(t,T)S_t + \beta(t,T) \]

4.3.3 Extended stream-flow prediction method

Extended stream-flow prediction (ESP) is a method for deriving runoff forecasts based on a spatially lumped precipitation-runoff model. The model is run up to the time of forecast with observed meteorological inputs. One major advantage of extended stream-flow prediction is the ability to conduct alternative scenario analyses. And the method may be better than index-variable methods in extreme years since the nonlinearities in the precipitation-runoff process are incorporated in the conceptual simulation model on which the procedure is based.

4.3.4 Time-series forecasting methods

Time series model can be used to describe the stochastic structure of the time sequence of stream flows and precipitation. Time series models are usually of the following form:

\[ X_t = \alpha_0 + \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q} \]

Where \( X_t = x_t - \mu_t \), \( x_t \) is the observed flow, \( \mu_t \) is the mean flow at time \( t \), \( \varepsilon_t \) is a normally distributed error term, \( \alpha_i \) and \( \beta_i \) are parameters to be estimated.

4.3.5 Updating long-term forecasts
Long-term forecast accuracy can often be improved through implementation of simple forecast updating procedures. The easiest approach to forecast updating is simply to develop an error correction equation as follow:

$$Q_f(t,T) = M(P,C,Q) + e_f(t,T)$$

Where $Q_f(t,T)$ represents the updated forecast made at time $t$ for lead time $T$, $M(P,C,Q)$ is the model estimate of the stream flow in the forecast period, and $e_f(t,T)$ is the error adjustment.

### 4.4 Review on Distributed Physically Based Hydrological Models

The history of physically based modelling of runoff generation, no more than 30 years old, has been characterised by an enormously wide range of approaches and levels of complexity.

One of the major difficulties in understanding and quantifying runoff generation in river basins terms from the presence of spatial variability in topography, geology, soil type, vegetation, etc. and in water fluxes, such as rainfall, infiltration and evapotranspiration. The accelerated growth of remote sensing distributed data, including satellite reading of terrain topography, soil type, land use, vegetation coverage, geology, soil moisture status, and radar estimates of rainfall rates, brings enormous promise to hydrologic science and runoff forecasting in particular.

Such distributed data are rapidly becomes widespread and easily available over vast regions of the world. In particular, digital elevation maps (DEMS) are ready available for most of the U.S. territory, and other regions of the world. Weather radar provides high temporal and spatial resolution estimates of rainfall-rates and benefits rainfall prediction enormously through detection of storm-system development. Near-full coverage of the U.S. and western Europe by weather radar networks is under implementation, therefore, there is a challenge to develop new model formulation which are capable of making best use of the distributed data available.

Distributed model framework has been presented that incorporates a variety of distributed data characterising terrain morphology soil parameters, and rainfall inputs, in the form of grid squares. Terrain morphology is obtained from DEMs, soil parameters are obtained from digitised maps of soil type, and rainfall rate may be provided by radar readings or by a network of tele-metering rain gages. The model attempts to describe the various physical
processes of runoff generation on the hill slope that are currently accepted, namely saturation-excess runoff, infiltration-excess runoff, interflow, and return flow.

Models of runoff generation have been developed for a wide variety of purposes, from the “one-off” design of engineering structures and water supply systems to modern real-time models used continuously in river regulation schemes. They are also proving valuable for studying the potential impacts of changes in land-use or climate. The variety of uses and the rapid increase both in scientific understanding and in technical support, from data collection systems and computer technology, have produced an enormous range in levels of sophisication.
5 INTRODUCTION OF THE SOFTWARE IN THE NETHERLANDS

5.1 Duflow

5.1.1 General introduction of Duflow

The Duflow Modelling Studio (DMS) supplies a complete set of tools to quickly perform easy analysis. On the other hand, the product is also capable of performing complex, integral studies.

Duflow Modelling Studio consists of the following components:

- **Duflow water quantity and quality**: With this program one can perform unsteady flow computations in networks of open watercourses. Duflow is also useful in simulating the transportation of substances in free surface flow and more complex water quality processes.

- **RAM precipitation runoff model**: With RAM one can calculate the supply of rainfall to the surface flow. RAM calculates the losses and delays that occur before the precipitation has reached the surface flow.

- **Tewor**: One can check the sewer system discharge on the surface water. It provides tools to prove the quality of water in urban environment.

- **MoDuflow** (implemented in DMS version 3.4): This program simulates an integrated ground water and surface water problem by combining the ground water model Modflow and Duflow.

5.1.2 RAM model

RAM is built on the basis of the following theory. First, a description of the precipitation and runoff process will be given out so that various hydrologic processes can be distinguished. Precipitation and runoff process are generally described at a catchment area level. Within a catchment area, the relevant parameters may vary substantially (soil type, slope, land use, vegetation etc). A detailed physical description of the occurring process is difficult to be given. In the literature, the hydrologic cycle is generally described as a chain of processes (indicating the course of processes and quantifying the amounts by means of water balances).

The hydrologic cycle is a continuous process in which water is evaporated from ocean and through the atmosphere and river, then back to the ocean. Fig 5-1 shows the various hydrologic processes.
The precipitation and runoff processes are distinguished between paved surface and unpaved surface. Within paved surface, a further distinction can be made. That is as the follow: paved surface in a rural area, urban area and Greenhouse area. The precipitation on paved surface in a rural area is discharged immediately by means of surface, or it is discharged by means of ground water due to infiltration outside the paved area. Precipitation in urban area will partly fall on paved surface such as roads and buildings and partly on unpaved areas such as parks, grasslands. The precipitation on the paved surface will be drained through the sewer system. The precipitation on unpaved surface will be drained through the soil or the drainage system. The precipitation on greenhouse is discharged directly or through a water storage reservoir to surface water. Within unpaved surface, it can also be further distinguished into 3 sub processes: infiltration into the soil moisture (unsaturated area),
percolation into the ground water (saturated area) and ground water discharge into the drainage system.

A water balance will be set up so that a general insight into the hydrologic cycle can be gained. A water balance for a catchment has the following form:

\[ \text{Incoming terms} = \text{Outflow terms} + \text{Change in storage in the area} \]

When the time interval for which the water balance is set up is long enough, the change in the storage of water in the area may be neglected. Shorter time intervals need to determine the runoff, because the total change in the storage of water in the area plays an important role in the relation of precipitation and runoff. The water balance is set up on the basis of fig 4.1. The incoming terms are the precipitation and water inlets and the outflow terms are the evaporation and the water outlets. The total storage in a catchment area is made up of surface storage, open water storage and storage in the unsaturated zone and the saturated zone. The above description is within a catchment level. The water balance can be further distinguished into two types: One is that of the unsaturated zone, the other is that of the saturated zone. Within the unsaturated zone, the incoming terms are infiltration, capillary rise from the saturated zone and lateral inflow. The outflow terms are evapotranspiration, interflow and percolation to the saturated zone. The storage term is formed by storage in unsaturated zone. The water balance function is as follow:

\[ Q_{\text{inf}} + Q_{\text{capillary}} + Q_{\text{li}} = Q_{\text{percolation}} + E + Q_{\text{interflow}} + \Delta_{\text{unsaturated}} \]

In which

\( Q_{\text{inf}} \) infiltration in mm
\( Q_{\text{capillary}} \) Capillary rise in mm
\( Q_{\text{li}} \) Lateral inflow soil moisture in mm
\( Q_{\text{percolation}} \) Percolation in mm
\( Q_{\text{interflow}} \) interflow in mm
\( \Delta_{\text{unsaturated}} \) storage change in unsaturated zone for the reflected time interval in mm

The water balance in saturated zones consists of the incoming terms: Percolation and lateral inflow and the outflow terms: Capillary to the unsaturated zone, seepage. The storage term is formed by the storage change in the ground water. The function is as follow:

\[ Q_{\text{percolation}} + Q_{\text{li}} = Q_{\text{capillary}} + K + W + \Delta_{\text{unsaturated}} \]

In which

\( \Delta_{\text{unsaturated}} \) storage change in saturated zone for the reflected time interval in mm

Some terms of the water balance in RAM are input and others such as evapotranspiration and ground water discharge are calculated by setting parameter. The water balance is important for calibration of precipitation runoff models. A precipitation runoff model is only capable of giving a correct description of the precipitation runoff processes when the water balance is described correctly.
Fig. 5-2 The sketch of the precipitation runoff module

Fig 5-2 shows the framework of the precipitation runoff model. The module is a simplified description of the above theory.

Within the open water surface, the effective precipitation can be simply calculated as precipitation minus the open water evaporation.

Within the paved surface, it is further divided into 3 types: urban area, rural area and greenhouse. In urban area, the precipitation is usually discharged through the sewer system. The way the precipitation is discharged is dependent on the sewer system. When extreme storm occurs, part of the precipitation will be discharged into the surface water by means of overflow. Within the urban area, a percentage of the surface can be defined to be open. In such a case, a part of the water will infiltrate into the unsaturated zone. The water infiltrated into the unsaturated zone is treated by the unpaved area part. In rural area, the precipitation is discharged directly through the drainage system. Same to the urban area, a percentage of open surface also can be defined. In greenhouse area, the water needed by greenhouse area can be supplied from a reservoir or a water company.

Within the unpaved surface, 3 processes can be further distinguished. Infiltration into the soil moisture in unsaturated zone, percolation into the ground water in saturated zone and ground water discharged into the drainage system. Within unsaturated zone, in surface depressions, a water balance is calculated for the precipitation. Part of it will infiltrate and part of it will be discharged as surface runoff. The amount of precipitation is decided by the infiltration capacity of soil. Where the infiltration capacity is assumed constant in time. If the precipitation intensity surpasses the infiltration capacity, the remaining part of the precipitation will be stored in the surface depression. When the storage in surface depression exceeds the maximum storage, the extra part will be discharged as surface runoff. Within saturated zone, a water balance of the amount of soil moisture is maintained in the soil moisture reservoir. The input item is the infiltration from the unsaturated zone and the output.
items are evapotranspiration and percolation. Both of the output items are decided by the soil moisture content. The actual value is described by means of a linear relation between the each item and the actual soil moisture. Finally the effective precipitation will discharge as ground water into the drainage system. Due to the resistance of the soil, there exists a significant slowing down process with a delayed discharge. This process is described by means of linear reservoir models. Besides the interflow and the drainage discharge, a slow and quick component is distinguished in the ground water discharge in the theoretical system. Two options can be chosen. One is method of two parallel Nash-cascades, the other is method of combination of Nash-cascade and Krayenhoff van de L X.eur. In the first option, both the quick and slow components are described by a number of linear reservoir in series, a Nash-cascade. In the second option, the hydrograph of the quick component is simulated by parallel linear reservoir while the slow component is described by a number of linear reservoirs in series.

![Fig. 5-3 Detailed processes in precipitation runoff module](image)

According to the simplified module, a detailed figure is given in fig5-3 to describe the various processes of the model.

**5.1.2 Application of Duflow**

A simple study case was given. The data adopted is the evaporation and precipitation of the reach up from Tangnaihai. Because of the limitation of the data collected, the parameter used is set arbitrarily and cannot represent the real situation of the reach. Fig5-4 and Fig 5-5 show the original data series and the simulated discharge. The detailed results are shown in
Table 5-1.

Table 5-1. Monthly curve of Evaporation and Precipitation

![Evaporation and Precipitation Curve](image)

Fig. 5-4 The curve of monthly Evaporation and Precipitation at Tangnaihai station

The curve of simulated monthly discharge

![Simulated Discharge Curve](image)

Fig. 5-5 The curve of simulated monthly discharge at Tangnaihai station
Tab. 5-1 Simulated result with the Duflow at Tangnaihai station

<table>
<thead>
<tr>
<th>Month</th>
<th>Evaporation</th>
<th>Precipitation</th>
<th>Simulated monthly discharge</th>
<th>observed discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.6</td>
<td>7.08</td>
<td>38.46</td>
<td>194</td>
</tr>
<tr>
<td>2</td>
<td>43.4</td>
<td>8.61</td>
<td>103.51</td>
<td>185</td>
</tr>
<tr>
<td>3</td>
<td>84.4</td>
<td>5.06</td>
<td>45.13</td>
<td>240</td>
</tr>
<tr>
<td>4</td>
<td>99.2</td>
<td>14.89</td>
<td>78.67</td>
<td>405</td>
</tr>
<tr>
<td>5</td>
<td>95.7</td>
<td>68.72</td>
<td>923.44</td>
<td>562</td>
</tr>
<tr>
<td>6</td>
<td>85.3</td>
<td>92.32</td>
<td>1962.99</td>
<td>588</td>
</tr>
<tr>
<td>7</td>
<td>97.6</td>
<td>162.45</td>
<td>3520.39</td>
<td>1800</td>
</tr>
<tr>
<td>8</td>
<td>104</td>
<td>100.67</td>
<td>3678.54</td>
<td>1310</td>
</tr>
<tr>
<td>9</td>
<td>75.6</td>
<td>112.55</td>
<td>3693.02</td>
<td>1360</td>
</tr>
<tr>
<td>10</td>
<td>75.7</td>
<td>32.31</td>
<td>2659.27</td>
<td>1170</td>
</tr>
<tr>
<td>11</td>
<td>52.9</td>
<td>4.39</td>
<td>1917.35</td>
<td>541</td>
</tr>
<tr>
<td>12</td>
<td>40.5</td>
<td>1.87</td>
<td></td>
<td>258</td>
</tr>
</tbody>
</table>

Without suitable parameters, the Duflow cannot be validated at present. However, according to the structure, several factors should be considered when it is used in runoff forecasting. The snowmelt runoff isn't concluded in the module, so when it is used in a situation that snowmelt runoff has a considerable influence, some methods should be used to estimate the snowmelt runoff. According to this model, it's simplified to make it easy to do parameter calibration. Therefore, some parameters do not have exact physical meaning. It's better to be calibrated with observed data.
5.2 PcRaster

PCRASTER is a software that can use GIS for the analysis and synthesis of spatial data in physical geography, soil science, hydrology and environment science. The software package, version 2 (1996) was developed at the Utrecht university by W.P.A. Van Deursen and C.G. Wesseling.

5.2.1 Introduction of Water Balance Model Based on Mechanism of Saturate Excess Runoff Generation in PcRaster Software Package

A very simple water balance model based on mechanism of saturation excess runoff generation was included in PCRASTER software package. And discharge process can be modelled with the water balance model. Precipitation is input of the model. First, evapotranspiration is subtracted from precipitation, the rest part called precipitation surplus. It is added to the deficient of soil water, a part of saturate excess can be rapid-runoff and percolation, the latter is added to groundwater, a part of groundwater drained to outflow point call delay-runoff. The sketch of water cycle was as follow:

Fig.5-6 sketch of water cycle

1. Evapotranspiration calculation
   \[ Evap = K \times Evap_{Ref} \]
   Where EvapRef is the reference crop evapotranspiration, K is the crop coefficient.

2. Precipitation surplus calculation
   \[ \text{Precipitation surplus} = \text{precipitation} - \text{Evap} \]

3. Soil water and its surplus calculation
   If precipitation surplus is positive, then it is added to the soil water.
If the maximum soil water content is reached, then
Soil water surplus = soil water + precipitation surplus – maximum soil water content

If precipitation surplus is negative, then soil water amount decreases by the absolute value of precipitation surplus on the condition of that some kind of minimum soil water content is not reached. And
Soil water surplus = 0
Anyway
Soil water >= minimum soil water content

(4) Rapid runoff and percolation calculation

For each interval, it is assumed that the soil water surplus is separated into two parts, one is rapid runoff, the other is percolation that is added to ground water. And the rapid runoff can drain to the outflow point in the same interval.
Rapid runoff = KD \times \text{soil water surplus}
Percolation = (1-KD) \times \text{soil water surplus}
Where KD is soil water surplus assignment coefficient.

(5) Delay runoff calculation

For each interval, a part of water in ground water drains to the outflow point, and cause a decrease in the amount of water in the groundwater storage.
Delay runoff = \text{groundwater}/C
Where C is a recession parameter, and depends on the geo-hydrological properties of a catchment

Obviously, the model has four parameters, they are the crop coefficient K, maximum soil water content, soil water surplus assignment coefficient KD and recession parameter C.

The water balance model is used to calculate water generation in each cell. And in the study area, runoff routing can be done with the local direction map, which has been made on the basis of digital elevation model (DEM).

5.2.2 Application of PcRaster Software for Discharge Simulation

In PCRASTER software package, a case study was given. Monthly area precipitation, evaporation and discharge process from November of 1993 to October of 1996 was plotted in the fig.5-7.

To simulate discharge process, recession parameter C was calibrated in the software. And former studies show that C value depends on the geo-hydrological properties of a catchment, and it has a lumped value between 4 and 7. Calibrated result shows the proper parameter value, 7. Other parameters in the model were determined according to actual circumstance or assumed. Crop coefficient differs from different land-use types. Its values were given as 0.5, 0.6 and 0.8 for pine, deciduous and mixed woods. Maximum soil water content and soil water surplus assignment coefficient were assumed to be 400mm and 0.2.
Fig. 5-7 Rainfall, evaporation and discharge process from Nov. 1993 to Oct. 1996

Fig. 5-8 Compare of observed and simulated discharge for the model calibration

Monthly-simulated discharge values were plotted against the recorded values and their distribution about the 1:1 line is shown in fig. 5-8. It shows that simulated values were very close to measured values and all points distributed uniformly about the 1:1 line. The result indicated the simple water balance model in PCRASTER software package has good simulated result for the case study area included in the PCRASTER package.
5.3 Caterpillar

The majority of statistical methods for analyzing hydrological data are based on the assumption that the hydrological processes are quasi-stationary. However, with the observed changes in climate this assumption is not quite correct anymore. It is known that there are many causes for the climate change. A number of causes follow from global processes such as sun activity, earth axis fluctuation, changes of earth zones, etc. A way for determination of these causes in the available datasets is by using the statistical Method of Main Components (MMC). This method is closely related with the Fourier method for spectral data analysis.

MMC was developed at the St.-Petersburg State University. It is an efficient statistical procedure and suitable software (Caterpillar®).

The basic algorithm of MMC consists of 4 parts. Each part will be considered in detail.

5.3.1 Reconstruction of 1D-series into 2D-series

Consider a time series \( (x_1, x_2, x_3, \ldots, x_M, \ldots, x_N) \) (1)

Let \( N \) be its length. Choose a number \( M \), which is known as the lag, \( (M < N) \). The larger the lag, the longer the period of the observed latent natural reasons fluctuations, and the smaller the number of the main components. Rewrite 1D-series \( (x_j) \) into 2D-series \( (x_{ij})_{i=1, j=1}^{k, j=m} \) as a matrix \( X \):

\[
X = (x_{ij})_{i=1, j=1}^{k, j=m} = \begin{pmatrix}
  x_1 & x_2 & x_3 & \ldots & x_M \\
  x_2 & x_3 & x_4 & \ldots & x_{k-1} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  x_k & x_{k+1} & x_{k+2} & \ldots & x_N \\
\end{pmatrix}
\]

(2)

This matrix may be shown as the MD (multi-dimensional) time series (which has volume \( k \)). Every column corresponds to the non-straight line consisting of \( k-1 \) parts:

\[
(x_1; x_2), (x_2; x_3), \ldots, (x_{k-1}; x_k)
\]

(3)

The same procedure can be done for the rows, however the consideration of columns is more suitable for theoretical explanation.

5.3.2 Analysis of the main components: singular separation of selected correlation matrix

Let us start from the calculation of the average values of columns:
and the standard deviation:

$$S_j = \sqrt{\frac{1}{k} \sum_{i=1}^{k} (x_{i,j-1} - \bar{x}_j)^2}$$

(5)

The found statistics, $\bar{x}_j$ and $S_j$, are to be used in the transformation of an initial series. The main components (MC) will be determined from the centered matrix $X^* = (x^*_i)$:

$$x^*_j = \frac{x_{i,j} - \bar{x}_j}{S_j}, i = 1, 2, \ldots, k; j = 1, 2, \ldots, M.$$ 

(6)

Eqn. (6) expresses the determination of centered values for every column. Also the determination of centered values for every column and for every row simultaneously can be performed. However, once centered values permit to leave an appearance of casual trends. Twice centered values lead to disappearance of information about the average trend. Thus, to make one procedure of the centering is recommended. As every column is a vector, the matrix $X$ may be seen as a series of $M$-dimensional vectors. Now the matrix $R$ is to be calculated:

$$R = \frac{1}{k} X^* (X^*)^T$$

(7)

$R$ is the correlation matrix, which consists of the elements $r_{ij}$:

$$r_{ij} = \frac{1}{k} \frac{1}{S_i S_j} \left( \frac{1}{k} \sum_{i=1}^{k} \frac{x_{i,j-1} - \bar{x}_j}{S_j} \right) \left( \frac{1}{k} \sum_{i=1}^{k} \frac{x_{i,j-1} - \bar{x}_i}{S_i} \right)$$

(8)

Note, that in the space of columns every element $r_{ij}$ is the cosinus of the angle between centered $k$-dimension vectors. The next part in the second step is the calculation of eigenvalues and eigenvectors of the matrix $R$, i.e. its decomposition, as follows:

$$R = P \Lambda P^T,$$

(9)

where $\Lambda$ is the diagonal matrix of eigenvalues:

$$\Lambda = \begin{pmatrix}
\lambda_1 & 0 & 0 & \ldots & 0 \\
0 & \lambda_2 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & 0 & \ldots & \lambda_M
\end{pmatrix}$$

(10)

and $P$ is the orthogonal matrix of eigenvectors of the matrix $R$: 
The matrix $P$ can be written as a transfer matrix of the main components $y_j$:

$$ X^* P = Y = (y_1, y_2, ..., y_M). $$

There is a very useful interpretation of the eigenvalues $\lambda_i$. The $\sqrt{\lambda_i}$-values are the selected standards of the corresponding main components $y_j$. It is proportional to the lengths of the ellipsoid described by the matrix $R$. Thus, $\sum_{j=1}^{M} \lambda_j = M$.

### 5.3.3 Selection of the main components

As the main components are orthonormal ($Y^*Y^* = I_M$), the initial series is to be decomposed into natural orthogonal components. Every vector $y_j$ may be considered as the corresponding eigenvector $p_j$ (because the vector $y_j$ is the result of the projection of the initial M-dimension series into the direction generated by $p_j$). In the same time the procedure $y_j = X^*p_j$ is the analogon of the linear reconstruction of the initial process by the roll operator. The MMC bores a number of filters, which is self-tuning in the components of the initial series.

Attention should be paid to the Nyquist frequency appearing in the case of too few measurements of the natural processes. Too few measurements lead to the appearance of false harmonics what do not have a natural prototype.

### 5.3.4 Reconstruction of 1D-initial series

The procedure of reconstruction is the key part of the MMC. It is founded by simple relationships. As it follows from Eqn. (12), the matrix $X^*$ may be reconstructed by the multiplication of $Y$ and $P^T$:

$$ X^* = YP^T = (y_1, y_2, ..., y_M) \cdot \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{pmatrix} = \sum_{i=1}^{M} y_i P_i^T = \sum_{i=1}^{M} X_i^* $$

Furthermore,
\[ X = x_k^M + X_s^M = X_0 + \sum_{i=1}^{M} X_i^M S = \sum_{i=0}^{M} X_i^M S \]  

We obtain the decomposition of the initial series by the sum of (M+1) series.

### 5.4 BestFit and Its Application

BestFit is a software provided by TU. The goal of BestFit is to find the distribution that best fits the input data. BestFit does not produce an absolute answer, it just identifies the distribution that is most likely to have produced the data. For a given distribution, BestFit varies the parameters of the function in order to optimize the goodness-of-fit. The results of a best fit calculation are only a best estimate, as it's nearly impossible to find an exact distribution to fit the data. The BestFit results are always evaluated quantitatively and qualitatively. Examining both the comparison graphs and statistics is prior to using a result.

There are 18 functions available in BestFit: Beta, Binomial, Chi-Square, Error Function, Erlang, Exponential, Gamma, Geometric, Hypergeometric, Logistic, Lognormal, Lognormal2, Negative Binomial, Normal, Pareto, Poisson, Triangular and Weibull.

The data of discharge of Tangnaihai station was inputted in Bestfit, and the input data is converted to a density distribution. A first guess of parameters is made using maximum-likelihood estimators, and the fit is optimized using the Levenberg-Marquardt method. The goodness-of-fit is measured for the optimized function. All functions are compared and the one with the lowest goodness-of-fit value is considered the best fit. Table and Graphs, Fig5-9, Fig5-10 and Fig5-11, are generated to show the results.

![Comparison of Input Distribution and Gamma](image)  

**Fig 5-9 Comparison of input distribution and Gamma (Tannaihai station)**
Comparison of Input Distribution and Gamma(1.77,3.73e+2)

Fig. 5-10 Comparison of input distribution and Gamma (Tannaihai station)

Difference Between Input Distribution and Gamma(1.77,3.73e+2)

Fig. 5-11 Difference between input distribution and Gamma (Tannaihai station)

Table 5-2 The Bestfit results (Tannaihai station)

<table>
<thead>
<tr>
<th>Minimum</th>
<th>Maximum</th>
<th>Mode</th>
<th>Mean</th>
<th>Std Deviation</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Input Settings:</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td>3550</td>
<td>283.9</td>
<td>661.5</td>
<td>496.8</td>
<td>2.47E+05</td>
<td>1.67</td>
<td>6.13</td>
<td>Type of Fit: Full Optimization</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Tests Run: Chi-Square K-S Test A-D Test</td>
</tr>
</tbody>
</table>

49
<table>
<thead>
<tr>
<th>Function</th>
<th>Chi-Square</th>
<th>Rank</th>
<th>K-S Test</th>
<th>Rank</th>
<th>A-D Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma(1.77,3.73e+2)</td>
<td>1.69E-04</td>
<td>1</td>
<td>0.178962</td>
<td>6</td>
<td>N/A</td>
</tr>
<tr>
<td>Lognormal2(6.26,0.66)</td>
<td>1.82E-04</td>
<td>2</td>
<td>0.145696</td>
<td>1</td>
<td>N/A</td>
</tr>
<tr>
<td>Lognormal(6.51e+2,4.77e+2)</td>
<td>1.82E-04</td>
<td>3</td>
<td>0.145696</td>
<td>2</td>
<td>N/A</td>
</tr>
<tr>
<td>Weibull(1.27,6.94e+2)</td>
<td>1.90E-04</td>
<td>4</td>
<td>0.155699</td>
<td>4</td>
<td>N/A</td>
</tr>
<tr>
<td>NegBin(1.00,1.51e-3)</td>
<td>2.77E-04</td>
<td>5</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Erlang(1.00,6.62e+2)</td>
<td>2.77E-04</td>
<td>6</td>
<td>0.2286</td>
<td>7</td>
<td>N/A</td>
</tr>
<tr>
<td>Expon(6.62e+2)</td>
<td>2.77E-04</td>
<td>7</td>
<td>0.230095</td>
<td>9</td>
<td>N/A</td>
</tr>
<tr>
<td>Geomet(1.51e-3)</td>
<td>2.79E-04</td>
<td>8</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Logistic(6.62e+2,2.72e+2)</td>
<td>1.65E-03</td>
<td>9</td>
<td>0.154012</td>
<td>3</td>
<td>N/A</td>
</tr>
<tr>
<td>Triang(1.12e+2,1.12e+2,3.55e+3)</td>
<td>2.10E-03</td>
<td>10</td>
<td>0.229622</td>
<td>8</td>
<td>N/A</td>
</tr>
<tr>
<td>Pareto(1.75,1.12e+2)</td>
<td>0.01009</td>
<td>11</td>
<td>0.685102</td>
<td>12</td>
<td>N/A</td>
</tr>
<tr>
<td>Normal(6.62e+2,4.97e+2)</td>
<td>0.11031</td>
<td>12</td>
<td>0.1608</td>
<td>5</td>
<td>N/A</td>
</tr>
<tr>
<td>Erf(1.39e-3)</td>
<td>34.945034</td>
<td>13</td>
<td>0.611536</td>
<td>11</td>
<td>N/A</td>
</tr>
<tr>
<td>Beta(22.88,27.36) * 3.44e+3 + 1.12e+2</td>
<td>4.31E+12</td>
<td>14</td>
<td>0.699014</td>
<td>13</td>
<td>N/A</td>
</tr>
<tr>
<td>Poisson(6.62e+2)</td>
<td>1.00E+34</td>
<td>15</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>HyperGeo(3.55e+3,3.55e+3,1.07e+4)</td>
<td>1.00E+34</td>
<td>16</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Chisq(6.61e+2)</td>
<td>1.00E+34</td>
<td>17</td>
<td>0.332878</td>
<td>10</td>
<td>N/A</td>
</tr>
<tr>
<td>Binomial(3.55e+3,0.19)</td>
<td>1.00E+34</td>
<td>18</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
</tr>
</tbody>
</table>

The results show that the gamma distribution is the best fitting for the monthly discharge of Tannaihai station.
6 APPLICATION OF STATISTIC METHODS FOR MID- AND LONG-TERM FLOW FORECASTING

The basic algorithm of the MMC is introduced in last Chapter. Caterpillar software can separate several main components of casual processes. The relative analysis reports and Caterpillar’s software was supplied by Dr. Pieter Van Gelder.

In this application we have the objects about month average discharge from January to June and water volume of year. The software we used is Caterpillar program. The results we obtained are data files and figures related with different parameters. The detail procedure was described as following.

6.1 Case Study 1: Application of MMC to Month Average Discharge from January to June

The month average discharge data series from January to June are analyzed and calculated with Caterpillar software. The suitable parameters are selected for every data series. 44 years (1956-1999) data are chosen for statistic and 2 years (2000-2001) data for forecasting in 46 years data. Generally the lag M is selected 18 and the principal components include first 6 main components. The detail introduction is following, January data for example.

![Fig. 6-1 Eigenfunctions of Tangnaihai data in January](image-url)
Fig. 6-2. The Principal Components of Tangnaihai Data in January

Fig. 6-3 The Reconstructed Series and Residuals of Tangnaihai data in January
Figure 6-4 and 6-5 show separately the absolute errors and relative errors of reconstruction of discharge series in January at Tangnaihai station. They indicate the average relative error is 9% with Maximum 39% (1960) and Minimum 0% (1987). There are five plots where the relative errors are more than 20%.

![Absolute Error of Reconstruction](image1.png)

**Fig. 6-4 Absolute Errors of Reconstruction Discharge Series in January at Tangnaihai**

![Relative Error of Reconstruction](image2.png)

**Fig. 6-5 Relative Errors of Reconstruction Discharge Series in January at Tangnaihai**

We forecasted the month discharge of two years in January with 2000 and 2001 for experiment. The results show the difference between forecasted and observed reached 42 m$^3$/s and 49 m$^3$/s separately with the relative errors 25% and 29%. It is not a satisfactory result.

The comparison of forecasted and observed for month average discharge show in table 6-1. Two half of year month discharge is forecasted for experiment. The average relative error is 31% with maximum of 51% in May of 2001 and minimum of 6% in June of 2000. However, if we forecasted with the multi-year average value of every month, the average relative error
is only 16%. So the forecast error with MMC is larger than the error with multi-year average value. It means the MMC is not efficient for month average discharge forecasting at Tangnaihai station. We can see this result more clearly in the figure 6-6 because the spots of multi-year average are closer to the relevantly observed spots.

Table 6-1. Comparison of forecast and observed for month average discharge

<table>
<thead>
<tr>
<th>year.month</th>
<th>observed</th>
<th>Forecast</th>
<th>Abs Error</th>
<th>Relative Error(%)</th>
<th>Multi- year Average</th>
<th>Abs error if forecast average</th>
<th>Relative Error(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000.01</td>
<td>167</td>
<td>209</td>
<td>42</td>
<td>25</td>
<td>170</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2000.02</td>
<td>156</td>
<td>217</td>
<td>61</td>
<td>39</td>
<td>167</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>2000.03</td>
<td>188</td>
<td>232</td>
<td>44</td>
<td>23</td>
<td>220</td>
<td>32</td>
<td>17</td>
</tr>
<tr>
<td>2000.04</td>
<td>380</td>
<td>478</td>
<td>98</td>
<td>26</td>
<td>363</td>
<td>-17</td>
<td>-4</td>
</tr>
<tr>
<td>2000.05</td>
<td>451</td>
<td>600</td>
<td>149</td>
<td>33</td>
<td>589</td>
<td>138</td>
<td>31</td>
</tr>
<tr>
<td>2000.06</td>
<td>936</td>
<td>877</td>
<td>59</td>
<td>6</td>
<td>900</td>
<td>-36</td>
<td>-4</td>
</tr>
<tr>
<td>2001.01</td>
<td>130</td>
<td>179</td>
<td>49</td>
<td>38</td>
<td>170</td>
<td>40</td>
<td>31</td>
</tr>
<tr>
<td>2001.02</td>
<td>136</td>
<td>175</td>
<td>39</td>
<td>29</td>
<td>167</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>2001.03</td>
<td>155</td>
<td>227</td>
<td>72</td>
<td>46</td>
<td>220</td>
<td>65</td>
<td>42</td>
</tr>
<tr>
<td>2001.04</td>
<td>312</td>
<td>417</td>
<td>105</td>
<td>34</td>
<td>363</td>
<td>51</td>
<td>16</td>
</tr>
<tr>
<td>2001.05</td>
<td>470</td>
<td>712</td>
<td>242</td>
<td>51</td>
<td>589</td>
<td>119</td>
<td>25</td>
</tr>
<tr>
<td>2001.06</td>
<td>856</td>
<td>1065</td>
<td>209</td>
<td>24</td>
<td>900</td>
<td>44</td>
<td>5</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
</tbody>
</table>

Fig. 6-6. The test curve of forecasted and observed month average discharge from January to June in 2000 and 2001
6.2 Case Study 2: Application of MMC to Forecast of Total Runoff of Year

We applied the MMC to forecast the total runoff of year in Tangnaihai station. The data series last 46 years from 1956 to 2001. The samples of 42 years is used for reconstruction and 4 years left for forecasting for experiment. The lag M of 18 is suitable and principal components are chosen as first 6 components. The procedure shows following. Figure 6-7 and 6-8 show separately the eigenfunctions and the principal components of Tangnaihai data of total runoff in year. The reconstructed series and its residuals are shown in figure 6-9.

![Eigenfunctions of Tangnaihai data in total runoff of year](image)

**Fig. 6-7 Eigenfunctions of Tangnaihai data in total runoff of year**
Fig. 6-8. The principal components of Tangnaihai data in total runoff of year

Fig. 6-9. The Reconstructed Series and Residuals of Tangnaihai data in year
The absolute and relative errors are shown in figure 6-10 and 6-11. Most years the relative errors of reconstruction are less than 20%. The average of relative error is 10% with Maximum of 41% in 1959 and Minimum of 0.6% in 1992.

![Absolute Error of Reconstructed series](image)

**Fig. 6-10. Absolute errors of reconstructed series of total runoff in year**

![Relative Error of Reconstructed Series](image)

**Fig. 6-11. Relative errors of reconstructed series of total runoff in year**

Figure 6-12 shows the forecast result of total runoff in year with MMC. The test forecast in last 4 years is closer to observed value. Only in 1999 the difference between forecasted and observed is large to 6.4 billion m$^3$ but the tendency of change is coherent between them. The detail results of forecast for experiment list in table 6-2. It shows the average relative error of 4 years is 13% with the maximum of 26% in 1999 and minimum of 2% in 1998. The multi-year average of total runoff is 20.4 billion m$^3$. The average relative error will be 27% if
we forecasted the total runoff in the 4 years with the multi-year average value. So we can say the MMC is available to forecast the total runoff of year in Tangnaihai station.

![Figure 6-12. The forecast results of total runoff in year](image)

### Table. 6-2. Comparison of Forecast and Observed of Total Runoff of Year

<table>
<thead>
<tr>
<th>year</th>
<th>Observed</th>
<th>Forecast</th>
<th>Difference</th>
<th>Relative Error</th>
<th>Multi-year Average</th>
<th>Abs error if forecast with average</th>
<th>Relative Error if forecast with average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>183</td>
<td>179</td>
<td>-4</td>
<td>2</td>
<td>204</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>1999</td>
<td>243</td>
<td>179</td>
<td>-64</td>
<td>26</td>
<td>204</td>
<td>-39</td>
<td>16</td>
</tr>
<tr>
<td>2000</td>
<td>152</td>
<td>170</td>
<td>18</td>
<td>12</td>
<td>204</td>
<td>52</td>
<td>34</td>
</tr>
<tr>
<td>2001</td>
<td>140</td>
<td>154</td>
<td>14</td>
<td>10</td>
<td>204</td>
<td>64</td>
<td>46</td>
</tr>
<tr>
<td>Average</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27</td>
<td></td>
</tr>
</tbody>
</table>

### 6.3 Conclusions about Statistic Model of MMC

Different parameters in the MMC have different influence on the results. So it is very important to select the parameters when you calculate. We have to test many parameters and then choose a suitable one for forecasting. The main parameter in the model is lag M. In the
application to Tangnaihai data the lag M are selected 18 and the main components is first 6 components.

All of the reconstruction results are better than the forecasted results. In two cases we applied the average relative errors of reconstructed series is 9% and 10% separately with January and year. But the forecasting experiment average error is 31% and 13% separately. The MMC is more efficient for forecasting of the year total runoff than one of the month average discharge in Tangnaihai station. The forecasting experiment shows that the forecasting relative error with MMC for month average discharge is larger than one with multi-year average. The forecasting relative error with MMC for total runoff of year is less than one with multi-year average.

6.4 Case Study 3: Application of MMC to Forecast the Monthly Discharge from July to December

The monthly data series from 1956 to 1985 are adopted. Altogether 6 data series are analyzed to forecast the monthly discharge of the last six months from 1986 to 1990 with Caterpillar software. Each data series, parameters are calibrated and optimised. The lag and the main components are varied from each other. Usually, it's better to adopt a bigger lag.

Fig 6-13 and Fig 6-18 show the monthly discharge forecasting situation from July to December and the detailed monthly discharge forecasting results of October to December is showed in table 6-3 to table 6-5.
Fig 6-14 Reconstructed and forecasted situation in August

Fig 6-15 Reconstructed and forecasted situation in September
Fig 6-16 Reconstructed and forecasted situation in October

Fig 6-17 reconstructed and forecasted situation in November
Fig 6-18 Reconstructed and forecasted situation in December

Table 6-3 Forecasting result in October

<table>
<thead>
<tr>
<th>YEAR (OCT)</th>
<th>INITIAL VALUE</th>
<th>FORECAST VALUE</th>
<th>ABSOLUTE ERROR</th>
<th>RELATIVE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>574</td>
<td>476</td>
<td>-98.4</td>
<td>-0.171</td>
</tr>
<tr>
<td>1987</td>
<td>532</td>
<td>684</td>
<td>152</td>
<td>0.285</td>
</tr>
<tr>
<td>1988</td>
<td>1220</td>
<td>834</td>
<td>-386</td>
<td>-0.317</td>
</tr>
<tr>
<td>1989</td>
<td>1150</td>
<td>1910</td>
<td>759</td>
<td>0.660</td>
</tr>
<tr>
<td>1990</td>
<td>681</td>
<td>2230</td>
<td>1549</td>
<td>2.276</td>
</tr>
</tbody>
</table>

Table 6-4 Forecasting result in November

<table>
<thead>
<tr>
<th>YEAR (NOV)</th>
<th>INITIAL VALUE</th>
<th>FORECAST VALUE</th>
<th>ABSOLUTE ERROR</th>
<th>RELATIVE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>332</td>
<td>291</td>
<td>-40.5</td>
<td>-0.122</td>
</tr>
<tr>
<td>1987</td>
<td>313</td>
<td>414</td>
<td>101</td>
<td>0.323</td>
</tr>
<tr>
<td>1988</td>
<td>611</td>
<td>640</td>
<td>28.8</td>
<td>0.047</td>
</tr>
<tr>
<td>1989</td>
<td>637</td>
<td>807</td>
<td>170</td>
<td>0.267</td>
</tr>
<tr>
<td>1990</td>
<td>400</td>
<td>713</td>
<td>313</td>
<td>0.782</td>
</tr>
</tbody>
</table>
Table 6-5 Forecasting result in December

<table>
<thead>
<tr>
<th>YEAR (DEC)</th>
<th>INITIAL VALUE</th>
<th>FORCAST VALUE</th>
<th>ABSOLUTE ERROR</th>
<th>RELATIVE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>203</td>
<td>180</td>
<td>-23.1</td>
<td>-0.114</td>
</tr>
<tr>
<td>1987</td>
<td>166</td>
<td>167</td>
<td>1.43</td>
<td>0.008</td>
</tr>
<tr>
<td>1988</td>
<td>256</td>
<td>216</td>
<td>-40.4</td>
<td>-0.158</td>
</tr>
<tr>
<td>1989</td>
<td>317</td>
<td>299</td>
<td>-17.8</td>
<td>-0.056</td>
</tr>
<tr>
<td>1990</td>
<td>213</td>
<td>341</td>
<td>128</td>
<td>0.602</td>
</tr>
</tbody>
</table>

Table 6-6 Comparison of the results between MMC and Multi-year Average methods

<table>
<thead>
<tr>
<th>MONTH</th>
<th>OBSERVED DISCHARGE</th>
<th>FORCASTED DISCHARGE (M1)</th>
<th>RELATIVE ERROR</th>
<th>MULTI-YEAR AVERAGE DISCHARGE</th>
<th>RELATIVE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>1700</td>
<td>1640</td>
<td>-0.035</td>
<td>1357</td>
<td>-0.202</td>
</tr>
<tr>
<td>8</td>
<td>699</td>
<td>611</td>
<td>-0.126</td>
<td>1136</td>
<td>0.625</td>
</tr>
<tr>
<td>9</td>
<td>1080</td>
<td>1910</td>
<td>0.769</td>
<td>1381</td>
<td>0.279</td>
</tr>
<tr>
<td>10</td>
<td>574</td>
<td>476</td>
<td>-0.172</td>
<td>1077</td>
<td>0.877</td>
</tr>
<tr>
<td>11</td>
<td>332</td>
<td>291</td>
<td>-0.122</td>
<td>506</td>
<td>0.525</td>
</tr>
<tr>
<td>12</td>
<td>203</td>
<td>180</td>
<td>-0.114</td>
<td>235</td>
<td>0.158</td>
</tr>
</tbody>
</table>

In order to obtain a generation realization of the forecasting results, a comparison with the results forecasted with multi-year average methods is shown in Table 6-6.

From the forecasting situations of each month, a conclusion is easy to be concluded. During the period July to October, only the changing trend is shown in the forecasting period, though the reconstructed series match the original one closely. During the period October to December, the reconstructed series also match the original one closely, and the forecasting result is satisfactory in the last three months, especially the first one or two years of the forecasting series. In 1986, the first forecasted year, the maximum absolute forecasting relative error is only 0.172 and the minimum is 0.114, while the maximum absolute forecasting relative error with multi-year average methods is 0.877 and the minimum is 0.158. So it can be adopted to apply in the low-flow season runoff forecasting. However, during the flood season, with more influential factions, it's hard to analyze and abstract the more decisive main components with quite short data series.
7 APPLICATION OF HYDROLOGICAL MODEL FOR MID- AND LONG-TERM FLOW FORECASTING

7.1 Xinanjiang Model and Its Application for Discharge Simulation

7.1.1 General introduction

The Xinanjiang model was developed in 1973 and published in 1980. Its main feature is the concept of runoff formation on repletion of storage, which means that runoff is not produced until the soil moisture content of the aeration zone reaches field capacity, and thereafter runoff equals the rainfall excess without further loss. This hypothesis was first proposed in China in the 1960s, and much subsequent experience support its validity for humid and semi-humid regions. According to the original formulation, runoff so generated was separated into two components using Horton's concept of final, constant, infiltration rate. Infiltrated water was assumed to go to the groundwater storage and the remainder to surface, or storm runoff. However, evidence of variability in the final infiltration rate. And in the unit hydrograph assumed to connect the storm runoff to the discharge from each sub-basin, suggested the necessity of a third component. Guided by the work of Kirkby(1978) an additional component. Interflow, was provided in the model in 1980. The modified model is now successfully and widely used in China.

7.1.2 The Structure of The Xinanjiang Model

The Concept of Runoff Formation on Repletion of Storage

Soil, or indeed any porous medium, possesses the ability of holding indefinitely against gravity a certain amount of water constituting storage. This is sometimes called “field moisture capacity”. By definition, the water held in this storage can’t become runoff and the storage can be depleted only by evaporation or the transpiration of the vegetation. Hence evaporation becomes the controlling soil moisture deficiency.

In humid regions, soil moisture can often reach field capacity within the entire soil cover. This implies that the replenishment in the next period of rainfall will be equal to the evaporation between rainfall and runoff viz. that before producing runoff the rainfall must satisfy the deficiency below field capacity which in turn is caused by the evaporation since field capacity last occurred and after that subsequent rainfall will all run off since soil has no more capacity to hold them. This is the concept of runoff formation on repletion of storage which is the basic concept of Xinanjiang model.

Outline of the Xinanjiang Model Structure
The basin is divided into a set of sub-basins. The outflow from each sub-basin is first simulated and then routed down the channels to the main basin outlet. Based on the concept of runoff formation on repletion of storage, the simulation of outflow from each sub-basin is consisted of four major parts:

(1) The evapotranspiration which generates the deficit of the soil storage, which is divided into three layers: upper, lower and deep;
(2) The runoff production which produces the runoff according the rainfall and soil storage deficit;
(3) The runoff separation which divided the above so determined runoff into three components: surface, subsurface and groundwater;
(4) The flow routing which transfers the local runoff to the outlet of each sub-basin forming the outflow of the sub-basin.

The flow chart of Xinanjiang model is show in Fig 7-1. All symbols inside the blocks are variables including inputs, outputs, state variables and internal variables while those outside are parameters.

![Flow chart for the Xinanjiang model](image)

Fig. 7-1 Flow chart for the Xinanjiang model

The inputs to the model are the areal mean rainfall, \( P \), and measured pan evaporation, \( EM \). The outputs are the discharge, \( TQ \). From the whole basin and the actual evapotranspiration, \( E \), which includes three components \( EU, EL, ED \).
The state variables are the areal mean tension water storage, $W$, and the areal mean free water storage, $S$. The areal mean tension water $W$ has three components $W_U$, $W_L$, and $W_D$ in the upper, lower and deep layer, respectively. The $FR$ is runoff contributing areal factor which is related to $W$. The rest of the symbols inside the blocks are all internal variables. $RB$ is the runoff directly from the small portion of impervious area. $R$ is the runoff produced from the pervious area and divided into three components $RS$, $RI$, and $RG$ referred to as surface runoff, interflow and groundwater runoff, respectively. The three components are further transferred into $QS$, $QI$, and $QG$ and together form the total inflow to the channel network of the sub-basin. The outflow of the sub-basin is $Q$.

**Parameters**

The characters outside the blocks are all parameters. $K$ is the ratio of potential evapotranspiration to pan evaporation if pan evaporation measurements are used as references. $WM$ and $B$ are two parameters describing the tension water distribution. $WM$ is the areal mean tension water capacity having components $UM$, $LM$ and $DM$. $B$ is the exponent of the tension water capacity distribution curve.

$IM$ is the factor of impervious area.

$SM$ and $EX$ are similar to $WM$ and $B$ while they describe the free water capacity distribution.

$KI$ and $KG$ are coefficients relating to $RI$ and $RG$.

$CI$, $CG$, $L$, $CS$, $KE$ and $XE$ are parameters for flow routing.

### 7.1.3 Application for Discharge Simulation

**Evaporation**

The potential evaporation

In Xinanjing model, the actual evaporation of the basin is related both to potential evapotranspiration and soil moisture condition. But direct determination of the potential evapotranspiration will encounter difficulties. For practical purpose, it is determined with the aide of the observed values from evaporation pan due to the fact that the data of pan evaporation to potential evapotranspiration is thus introduced into the model and calibrated mainly by the trial-and-error method.

**Three-Layer soil moisture model**
Traditionally, the actual evapotranspiration is treated to be proportional to the soil moisture. This is considered to be one layer soil model.

The main drawback of the one-layer computation lies in the fact that the distribution of moisture in the vertical profile cannot be taken into account, which result in considerable error in the actual evaporation calculation, particularly when light rain falls in a long period of drought. This is because that the antecedent soil moisture content is already very low, thus the calculated value of evaporation by one-layer will be very small. But in reality, the part of soil moisture replenished by new rainfall will be mainly distributed in the top soil and, in addition, if there is a good vegetative cover, a certain portion of the rain will be intercepted by the vegetation. The moisture which remains in the topsoil as well as on the branches and leaves of plants will evaporate at a rate comparable with the potential evapotranspiration and, therefore, can no longer be taken the same thing as the antecedent soil moisture content. So that a top-layer is necessary to be introduced into the soil model for treating the new rain different from the old soil moisture.

Meanwhile, during dry season, still another error in the calculation of actual evaporation may occur especially in humid and well covered area. In this case, though the topsoil is dried out, the deep layer may be still moist. The root system of vegetation will grow deeper to absorb the moisture there for its transpiration. Hence the rule of direct proportion no longer holds at this moment. It is necessary to further improve the calculation by putting forward a deeper layer to the model. Thus a three-layer soil moisture model was so gradually developed.

**Calculation Rules of evapotranspiration**

The total areal mean soil moisture capacity of a basin, WM, is divided into three portions: UM the upper part, LM the lower part and DM the deeper part. WU, WL and WD are storage at any moment corresponding to the three layers. Evaporation occurs at the potential rate, equal to K times the pan evaporation rate in the case of using pan evaporation measurements,

\[ EU = K \cdot EM \]

Until the storage WU of the uppermost layer is exhausted. On exhaustion of the upper layer, any remaining potential evapotranspiration is applied to the lower layer, but the efficiency is
modified by multiplication by the ratio of the actual storage WL to the capacity storage LM of
the layer.

\[ EL = (K \cdot EM - EU) \cdot WL / LM \]

When the lower layer storage WL is reduced to a specified proportion, C, of LM, evapotranspiration is assumed to continue, but at a rate ED given by

\[ ED = C \cdot (K \cdot EM - EU) \]

Runoff production

The concept of runoff formation on repletion of storage holds true so far as one single
point, in the sense of a very small area, of a basin is concerned. For an entire watershed, things are more complicate. The moisture deficit often varies from place to place. This non-uniform distribution effects the runoff production of a whole basin significantly. To solve the problem, in Xinanjiang model, a distribution of tension water capacity is suggested in statistical manner.

To provide for a non-uniform distribution of tension water capacity throughout the basin or sub-basin whatever is concerned, a tension water capacity curve is introduced, see Fig.

In Fig.7-2 f/F represents the proportion of the pervious area of the basin whose tension
water capacity is less than or equal to the value of the ordinate W'M. The tension water
capacity at a point, W'M, varies from zero to a maximum MM according to the relationship

\[ (1 - f / F) = (1 - W' M / MM)^p \cdot (1 - IM) \]

Where MM and B are parameters.

The areal mean tension water capacity, WM, constitutes an alternative parameter to the maximum value MM. These are related through the parameter B. It is easy show that

\[ MM = WM \cdot (1 + B) / (1 - IM) \]
The state of the basin, at any time, is assumed to be represented by a point $x$ is proportional to the areal mean tension water capacity.

![Diagram of tension water capacity distribution](image)

**Fig.7-2 The distribution of tension water capacity in the basin**

When rainfall exceeds evaporation, the ordinate of Fig.7-2 is increased by the excess, $x$ moves upwards along the curve and runoff is generated proportional to the area shown shaded to the left and above the point $x$ in Fig.7-2.

If $P - K \cdot EM + AU$ is less then $MM$, then

$$R = P - K \cdot EM - WM + W + WM \cdot \left(1 - \frac{(P - K \cdot EM + AU)}{MM}\right)^{(1+B)}$$

otherwise

$$R = P - K \cdot EM - WM + W$$

The runoff $R$, generated in a wet period in accordance with Fig7-2., should be further separated into its three components, RS surface runoff, RG ground water runoff, and RI interflow. To effect this, the concept of free water storage and its distribution are developed.

Distribution of free water capacity is assumed in Xinanjinag model that the free water storage and its capacity are non-uniformly distribute over the area, FR, on which the runoff is currently produced. The distribution curve is illustrated in Fig7-2.. The free water storage
capacity SM are assumed to be distributed between zero and a point maximum MS in a parabolic manner, over FR.

\[(1 - f / F) = (1 - SM / MS)^{Ex}\]

where f is the portion of the basin area for which the free water storage capacity is less than or equal to SM and Ex is a parameter.

![Fig.7-3 The separation of runoff components](image)

It is further assumed that the current state of free water storage in the basin can be represented by a point (ordinate BU on the parabola of Fig.7-3) implying that the portion of the basin to the left of that point is at capacity storage and to the right the storage is constant, below capacity level.

The areal mean free water storage capacity SM maybe used instead of MS as a parameter

\[SM = MS / (1 + Ex)\]

By integration of S'M and substitution of SM for MS, the equivalent free water storage S over the producing area FR can be found from the equation below

\[1 - \frac{S}{SM} = (1 - \frac{BU}{MS})^{(1 + Ex)}\]
So, the runoff $R$ generated in accordance with Fig. and expressed as the depth $P - K \cdot EM$ over the runoff producing area of the basin is applied by adding $P - K \cdot EM$ to BU in Fig., yielding a contribution $RS$ to surface runoff.

Algebraically, if $BU + P - K \cdot EM < MS$, then

$$RS = (P - K \cdot EM - SM + S + SM \cdot (1 - (P - K \cdot EM + BU) / MS)^{t+e}) \cdot FR$$

Otherwise,

$$RS = (P - K \cdot EM + S - SM) \cdot FR$$

The remainder of $R$ becomes an addition, $\Delta S$, to the free water storage $S$, which in turn contributes $RI$ laterally to inflow and $RG$ vertically to groundwater, according to the relations

$$RI = S \cdot KI \cdot FR$$
$$RG = S \cdot KG \cdot FR$$

where $KI$ and $KG$ are parameters.

Xinanjiang model mainly simulate the hourly or daily runoff, all parameters related to rates will change with the time scale that the model based on. As the time scaling up further to a monthly-base, the model structure may get very simple, it can be simplified a water balance model.

The application of Xinanjiang model in Yellow River basin will be finished in China after getting the digital map.
7.2 NAM Model and Its Application for Discharge Simulation

7.2.1 Introduction of NAM Model

The NAM hydrological model simulates the rainfall-runoff processes occurring at the catchment scale. NAM forms part of the rainfall runoff module of the MIKE11 River modelling system. And it can either be applied independently or used to represent one or more contributing catchments that generate lateral inflows to a river network. NAM is the abbreviation of Danish "Nedbor-Afstromnings-Model", meaning precipitation runoff model, and was originally developed by the department of hydrodynamics and water resources at the technical university of Denmark (Nielsen and Hansen, 1973). NAM model can be characterised as a deterministic, lumped, conceptual model with moderate input data requirements. The model has been applied to a number of catchments around the world, representing many different hydrological regimes and climatic conditions.

Input of NAM model includes four parts: model parameters, initial conditions, meteorological data and Stream flow data for model calibration and validation. And the basic meteorological data requirements are: Rainfall, temperature and Potential evapotranspiration.

![fig.7-4 structure of NAM model](image)

**Fig.7-4 structure of NAM model**

**Model structure**

NAM model treats each catchment as a single unite, and parameters and variables represent, therefore, average values for the entire catchment. As a result some of the model parameters can be evaluated from physical catchment data, but the final parameter estimation must be performed by calibration against time series of hydrological observations. The model structure is shown in figure 7-4. Nam simulates the rainfall runoff process by continuously accounting for the water content in three different and mutually interrelated storages that
represent different physical elements of the catchment. Those storages are: Surface storage, Low or root zone storage and Ground water storage

Basic modelling components

Surface storage

Moisture intercepted on the vegetation as well as water trapped in depression and in the uppermost, cultivated part of the ground is represented as surface storage. \( U_{\text{max}} \) denotes the upper limit of amount of water in the surface storage.

The amount of water, \( U \), in the surface storage is continuously diminished by evaporative consumption as well as by horizontal leakage (interflow). When there is maximum surface storage, some of the excess water, \( P_N \), will enter the stream as overland flow, whereas the remainder is diverted as infiltration into the lower zone and groundwater storage.

Lower zone or root zone storage

The soil moisture in the root zone, a soil layer below the surface from which the vegetation can draw water for transpiration, is represented as lower zone storage. \( L_{\text{max}} \) denotes the upper limit of the amount of water in this storage.

Moisture in the lower zone storage is subject to consumptive loss from transpiration. The moisture content controls the amount of water that enters the groundwater storage as recharge and the interflow and overland flow components.

Evapotranspiration

Evapotranspiration demands are first met at the potential rate from the surface storage. If the moisture content \( U \) in the surface storage is less than these requirements \( (U < E_p) \), the remaining fraction is assumed to be withdrawn by root activity from the lower zone storage at an actual rate \( E_a \). \( E_a \) is proportional to the potential evapotranspiration and varies linearly with the relative soil moisture content, \( L/L_{\text{max}} \), of the lower zone storage:

\[
E_a = (E_p - U) \cdot L/L_{\text{max}}
\]

Overland flow

When the surface storage spills, i.e. when \( U > U_{\text{max}} \), the excess water \( P_N \) gives rise to overland flow as well as to infiltration. \( QOF \) denotes the part of \( P_N \) that contributes to overland flow. It is assumed to be proportional to \( P_N \) and to vary linearly with the relative soil moisture content, \( L/L_{\text{max}} \), of the lower zone storage:

\[
QOF = C_{QOF} \cdot \frac{L/L_{\text{max}} - TOF}{1 - TOF} \quad P_N > TOF
\]

\[
0 \leq C_{QOF} \leq 1, \quad 0 \leq TOF \leq 1
\]

Where \( C_{QOF} \) is the overland flow runoff coefficient \( (0 \leq C_{QOF} \leq 1) \), and \( TOF \) is the threshold value for overland flow \( (0 \leq TOF \leq 1) \).

The proportion of the excess water \( P_N \) that does not run off as overland flow infiltrates into the zone storage. A portion, \( D_L \), of the water available for infiltration, \( (P_N - QOF) \), is assumed to increase the moisture content \( L \) in the lower zone storage. The remaining amount of infiltration moisture, \( G \), is assumed to percolate deeper and recharge the groundwater
storage.

Interflow

The interflow contribution, QIF, is assumed to be proportional to U and to vary linearly with the relative moisture content of the lower zone storage.

\[
QIF = \left(CKIF\right) \frac{L/L_{\text{max}} - TIF}{1 - TIF} U, \quad L/L_{\text{max}} > TIF \\
0, \quad \ldots \ldots, L/L_{\text{max}} \leq TIF
\]

where CKIF is the time constant for interflow, and TIF is the root zone threshold value for interflow \(0 \leq TIF \leq 1\)

**Interflow and overland flow routing**

The interflow is routed through two linear reservoirs in series with the same time constant CK12. The overland flow routing is also based on the linear reservoir concept but with a variable time constant

\[
CK = CK_{12} \left(OF < OF_{\text{min}}\right)^{\beta}, \ldots OF \geq OF_{\text{min}}
\]

Where OF is the overland flow (mm/hour), OF_{min} is the upper limit for linear routing (=0.4mm/hour), and \(\beta=0.4\)

The constant \(\beta=0.4\) corresponds to using the Manning formula for modelling the overland flow. Equation (4.4) ensures in practice that the routing of real surface flow is kinematic, while subsurface flow being interpreted by NAM as overland flow (in catchments with no real surface flow component) is routed as a linear reservoir.

Groundwater recharge

The amount of infiltrating water G recharging the groundwater storage depends on the soil moisture content in the root zone

\[
G = (PN - QOF) \frac{L/L_{\text{max}} - TG}{1 - TG}, \ldots L/L_{\text{max}} > TG \\
0, \ldots, L/L_{\text{max}} \leq TG
\]

Where TG is the root zone threshold value for groundwater recharge \(0 \leq TG \leq 1\)

Soil moisture content

The lower zone storage represents the water content within the root zone. After apportioning the net rainfall between overland flow and infiltration to groundwater, the remainder of the net rainfall increase the moisture content L within the lower zone storage by the amount DL

\[
DL = P_n - QOF - G
\]

Baseflow

The base flow BF from the groundwater storage is calculated as the outflow from a linear reservoir with time constant CKbf
7.2.2 Monthly Discharge Simulation

Input monthly average area rainfall, potential evaporation, air temperature and recorded discharge into NAM model, calibrate parameters and simulate discharge. Monthly recorded and simulated discharges are shown in Fig.3.5. Statistical value of result is shown in Table 3.1 and accumulative discharges are compared in Fig.3.6.

Fig.7-5 shows that simulated discharge is much well with recorded in most years though simulated peak discharge is higher or lower than the recorded in another years. And statistic results in Table 3.2 indicate simulation efficiency coefficient is nearly up to 0.65 and correlation coefficient between simulated and recorded discharge is also high (0.816). Multi-year average simulated annual discharge (172.2 mm) is very close to recorded one (170.9 mm). And accumulative value of recorded and simulated discharge (Fig.7-6) also shows water balance between recorded and calculated discharge. The figures and table illustrate simulation results by NAM is acceptable.

![Fig.7-5 recorded and simulated discharge at Tangnaihai station from 1956 to 1975](image)

![Fig.7-6 compared of accumulative value of recorded and simulated discharge](image)
Table 7-1 statistical value of simulation results by NAM model

<table>
<thead>
<tr>
<th>Annual (mm)</th>
<th>Recorded</th>
<th>Simulated</th>
<th>Relative error(%)</th>
<th>Efficency coefficient</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>170.9</td>
<td>172.2</td>
<td>0.8</td>
<td>0.649</td>
<td>0.816</td>
<td></td>
</tr>
</tbody>
</table>

Table 7-2 Comparison of recorded and calculated discharge by Nam model

<table>
<thead>
<tr>
<th></th>
<th>1974</th>
<th></th>
<th>1975</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>recorded</td>
<td>calculated</td>
<td>absolute error (mm)</td>
<td>relative error(%)</td>
</tr>
<tr>
<td>1</td>
<td>3.2</td>
<td>6.2</td>
<td>3.0</td>
<td>94.6</td>
</tr>
<tr>
<td>2</td>
<td>3.1</td>
<td>5.5</td>
<td>2.5</td>
<td>80.0</td>
</tr>
<tr>
<td>3</td>
<td>4.6</td>
<td>5.1</td>
<td>0.6</td>
<td>12.5</td>
</tr>
<tr>
<td>4</td>
<td>6.4</td>
<td>5.0</td>
<td>-1.4</td>
<td>-22.0</td>
</tr>
<tr>
<td>5</td>
<td>10.7</td>
<td>5.4</td>
<td>-5.4</td>
<td>-49.9</td>
</tr>
<tr>
<td>6</td>
<td>18.6</td>
<td>11.9</td>
<td>-6.7</td>
<td>-36.0</td>
</tr>
<tr>
<td>7</td>
<td>18.8</td>
<td>21.4</td>
<td>2.6</td>
<td>13.6</td>
</tr>
<tr>
<td>8</td>
<td>21.9</td>
<td>26.3</td>
<td>4.4</td>
<td>20.2</td>
</tr>
<tr>
<td>9</td>
<td>28.5</td>
<td>31.9</td>
<td>3.4</td>
<td>12.1</td>
</tr>
<tr>
<td>10</td>
<td>20.6</td>
<td>21.1</td>
<td>0.6</td>
<td>2.7</td>
</tr>
<tr>
<td>11</td>
<td>12.0</td>
<td>13.1</td>
<td>1.0</td>
<td>8.6</td>
</tr>
<tr>
<td>12</td>
<td>5.6</td>
<td>8.8</td>
<td>3.2</td>
<td>57.3</td>
</tr>
</tbody>
</table>

7.3 MWBM and Its Application for Discharge Simulation

7.3.1 Introduction of Monthly Water Balance Model (MWBM)

With the development of computer technique, Large scales hydrological model was widely used in assessment of environment change, flow forecasting and any other aspects. Considering the characteristics of runoff yielding in Yellow River, a monthly water balance model was established according to water equilibrium principle. Calculated discharge in the model consists of surface flow, ground flow and snow-melting flow. Flow sink was not considering as calculation interval is large enough and all runoff components can flow out of the basin. Model structure is shown in Fig 7-7.
Four parameters are in the model, and monthly average watershed precipitation, evaporation and air-temperature is model’s input. Monthly runoff and other mid-variables can be outputted. They are soil storage capacity $S_{\text{max}}$, surface flow calculation coefficient $k_s$, ground flow calculation coefficient $k_g$ and snow-melting runoff coefficient $k_m$ respectively. The model has simple structure, less parameter and larger application area compared with other hydrological models.

Separation between rainfall and snowfall is prerequisite of surface flow and snow melting flow calculation. Precipitation forms rely on air temperature and could be separated into rainfall and snowfall. According to the recorded data of precipitation and air temperature, the highest air temperature $T_H$ and the lowest $T_L$ are chosen first, they are 4 °C and -4 °C respectively. Rainfall is equal to precipitation when air temperature is above +4 °C, and snowfall is equal to precipitation when air temperature is below -4 °C. Snowfall can be estimated by linear interpolation when temperature occurred between $T_H$ and $T_L$. The estimation function as follow:

$$P_{SN} = \frac{T_H - T_I}{T_H - T_L} \cdot P_I$$

$$P_{RJ} = P_I - P_{SN}$$

Where $P_I$, $P_{RJ}$, $P_{SN}$ and $T_I$ are precipitation, rainfall, snowfall and air temperature in the calculation interval $I$ respectively.

The surface flow is directly proportional to soil moisture and precipitation. And surface
flow calculation formula can be written as:

\[ Q_{si} = k_s \cdot \frac{S_{i-1}}{S_{max}} \cdot P_{ri} \]

Where \( Q_{si} \) and \( P_{ri} \) is surface flow and rainfall in the calculation interval \( i \). \( S_{i-1} \) is soil moisture in the calculation interval \( i-1 \).

The ground flow is calculated by a linear reservoir. The equation is as follow.

\[ Q_{gi} = k_g \cdot S_{i-1} \]

Where \( Q_{gi} \) is ground flow in the calculation interval \( i \).

Rainfall, snowfall separation and snow accumulation estimation are prerequisite of snow melting flow calculation. And snow melting rate is not only an exponential function with air temperature, but also proportional to snow accumulation. So snow melting flow formula is as following.

\[ Q_{sn} = k_{sn} \cdot \frac{T_t - T_l}{T_n - T_l} \cdot SN_i \]

\[ SN_i = SN_{i-1} + P_{SN} \]

Where \( Q_{sn} \) and \( SN_i \) are snow melting flow and snow accumulation in the calculation interval \( i \). \( SN_{i-1} \) is snow accumulation in the calculation interval \( i-1 \).

Actual basin evaporation can be calculated by one layer soil evaporation model, the calculation formula as following

\[ E_i = \frac{S_{i-1}}{S_{max}} \cdot E_{601} \]

Where \( E_i \) is actual basin evaporation in the calculation interval \( i \). \( E_{601} \) is water surface evaporation measured by 601 evaporation ware in the calculation interval \( i \).

For land-surface system, Precipitation is input variable and output variables are surface flow and infiltration. For soil system, rainfall infiltration is input variable and soil evaporation; ground flow is output variables. According to water equilibrium principle, soil storage water can be calculated with the following equation.

\[ S_t = S_{i-1} + P_t - Q_{si} - E_i - Q_{gi} - Q_{SN} \]

Linear summing up surface flow, ground flow and snow melting flow consist of calculation runoff.

\[ Q_{ai} = Q_{si} + Q_{gi} + Q_{SN} \]

Where \( Q_{ai} \) is analogised discharge in the calculation interval \( i \).
7.3.2 Parameters Calibration and Discharge Simulation

Monthly average watershed precipitation, air temperature, evaporation ability and recorded runoff in the study area were used as input for the model. The calibration data is from 1960 to 1969 and validation period is from 1970 to 1975. Nash efficiency coefficient and annual mean relative error is set as objective function to calibrate model parameters. Recorded and simulated monthly discharge at Tangnaihai station from 1960 to 1975 are compared and shown in Fig.7-7. Simulation result was given in Table7-2.

Fig.7-8 Simulated and recorded discharge process at Tangnaihai station from 1956 to 1975
Figure 7-8 shows that the model has good simulation result to Dalihe basin. And monthly simulated discharge matches well with the observed discharge. Table 7-1 indicates annual simulated discharge in the calibration period and validation period is close to the recorded, simulation efficiency coefficient in the two periods is 68.8% and 61.1% respectively.

In order to study runoff components, process of simulated monthly surface flow, snowmelt flow and groundflow was plotted in Fig 7-9. And rate of each component is also listed in Table 7-3. Fig 7-9 shows that the three components contribute to discharge in flood period and ground flow is main runoff component in lower period.

Table 7-3 Statistic value of simulation result from MWBM

<table>
<thead>
<tr>
<th></th>
<th>Calibration period</th>
<th>Validation period</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Annual discharge(mm)</td>
<td>Efficiency coefficient(%)</td>
</tr>
<tr>
<td></td>
<td>recorded</td>
<td>simulated</td>
</tr>
<tr>
<td>Annual discharge</td>
<td>172.0</td>
<td>173.4</td>
</tr>
</tbody>
</table>
Fig. 7-10 Compare of annual recorded and simulated discharge

Fig. 7-11 Multi-year average monthly recorded and simulated discharge distribution in a year from 1960 to 1975

Annual simulated runoff values were plotted against the recorded value and their distribution about 1:1 line is shown in fig7-10. The figure indicated the simulation runoff value were slightly above the 1:1 line for lower values of recorded runoff. For high value of the measured, the simulated were slightly below the 1:1 line. Regression analysis was performed between the observed and simulated runoff and best-fit line is shown in fig7-11. The high value (0.91) of the correlative coefficient indicated a close relationship between measured and simulated runoff.

Multi-year average monthly recorded and simulated discharge distribution in a year from 1960 to 1975 given in fig7-11. The figure also indicated simulated discharge is closed to
recorded ones.

Table 7-4 components of simulated discharge at Tangnaihai station

<table>
<thead>
<tr>
<th></th>
<th>Surface flow</th>
<th>Ground flow</th>
<th>Snowmelt flow</th>
<th>Discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume (mm)</td>
<td>51.4</td>
<td>95.2</td>
<td>23.3</td>
<td>169.9</td>
</tr>
<tr>
<td>Percentage (%)</td>
<td>30.3</td>
<td>56.0</td>
<td>13.7</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 7-5 Comparison of recorded and calculated discharge by MWBM model

<table>
<thead>
<tr>
<th></th>
<th>1974</th>
<th></th>
<th>1975</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>recorded</td>
<td>calculated</td>
<td>absolute error (mm)</td>
<td>relative error(%)</td>
</tr>
<tr>
<td>1</td>
<td>3.2</td>
<td>4.0</td>
<td>0.8</td>
<td>24.9</td>
</tr>
<tr>
<td>2</td>
<td>3.1</td>
<td>3.2</td>
<td>0.1</td>
<td>2.6</td>
</tr>
<tr>
<td>3</td>
<td>4.6</td>
<td>2.3</td>
<td>-2.3</td>
<td>-49.5</td>
</tr>
<tr>
<td>4</td>
<td>6.4</td>
<td>7.4</td>
<td>1.0</td>
<td>16.2</td>
</tr>
<tr>
<td>5</td>
<td>10.7</td>
<td>13.0</td>
<td>2.2</td>
<td>20.9</td>
</tr>
<tr>
<td>6</td>
<td>18.6</td>
<td>30.4</td>
<td>11.8</td>
<td>63.4</td>
</tr>
<tr>
<td>7</td>
<td>18.8</td>
<td>21.0</td>
<td>2.1</td>
<td>11.4</td>
</tr>
<tr>
<td>8</td>
<td>21.9</td>
<td>22.0</td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>9</td>
<td>28.5</td>
<td>23.5</td>
<td>-4.9</td>
<td>-17.3</td>
</tr>
<tr>
<td>10</td>
<td>20.6</td>
<td>16.1</td>
<td>-4.4</td>
<td>-21.5</td>
</tr>
<tr>
<td>11</td>
<td>12.0</td>
<td>8.1</td>
<td>-3.9</td>
<td>-32.4</td>
</tr>
<tr>
<td>12</td>
<td>5.6</td>
<td>6.3</td>
<td>0.7</td>
<td>12.5</td>
</tr>
</tbody>
</table>
8 RESULTS AND CONCLUSIONS

Monthly discharge from 1986 to 1990 was forecasted with MWBM, NAM and MMC method respectively. Recorded and forecasted discharge in each year were compared and shown in fig8-1 to fig8-5. And forecasted errors were also calculated and listed in the table8-1.

Fig.8-1 Monthly recorded and forecasted discharge in 1986

Fig.8-2 Monthly recorded and forecasted discharge in 1987
Fig.8-3 Monthly recorded and forecasted discharge in 1988

Fig.8-4 Monthly recorded and forecasted discharge in 1989

Fig.8-5 Monthly recorded and forecasted discharge in 1990

Fig 8-1 to fig8-5 show that forecasted discharge by the three models are closed to the
recorded in most low flow months, but in some flood months, the forecasted value is much different with recorded one. And forecasted accuracy by MWBM nearly matches with that by NAM model.

Statistical results in table 8-1 indicate that the three models have big forecasted errors in 1989, average relative errors in the year are all excess 60%. In the other four years, MWBM and NAM models have similar forecasted errors; average relative forecasting errors fall in the range from 15% to 55%. For MMC method, average relative forecasted error in 1987 is in smallest value of 23.5% and that in 1990 is in highest value of 102.8%. Forecasted errors in the first two years are relative lower than that in the rest three years, it has increasing trendy.

If acceptable forecasted result is that relative error is less than 20% or absolute error is less than 4mm, then acceptable percentage of forecasted discharge was also calculated and listed in table 8-1. Result shows that most forecasted value by MWBM and NAM is qualified. For MMC method, only in the first two years and low flow period of other years, most forecasted value is acceptable.

<table>
<thead>
<tr>
<th>Year</th>
<th>Models</th>
<th>Absolute error (mm)</th>
<th>Relative error (%)</th>
<th>Acceptable percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Max</td>
<td>Min</td>
<td>Average</td>
</tr>
<tr>
<td>1986</td>
<td>MWBM</td>
<td>11.2</td>
<td>0.2</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>NAM</td>
<td>7.8</td>
<td>0.4</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>MMC</td>
<td>23.1</td>
<td>0.3</td>
<td>5.9</td>
</tr>
<tr>
<td>1987</td>
<td>MWBM</td>
<td>10.3</td>
<td>0.0</td>
<td>2.9</td>
</tr>
<tr>
<td></td>
<td>NAM</td>
<td>18.9</td>
<td>0.8</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>MMC</td>
<td>23.9</td>
<td>0.1</td>
<td>4.9</td>
</tr>
<tr>
<td>1988</td>
<td>MWBM</td>
<td>7.4</td>
<td>0.1</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>NAM</td>
<td>5.2</td>
<td>0.4</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>MMC</td>
<td>34.1</td>
<td>0.0</td>
<td>7.8</td>
</tr>
<tr>
<td>1989</td>
<td>MWBM</td>
<td>17.2</td>
<td>0.1</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>NAM</td>
<td>10.2</td>
<td>0.5</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>MMC</td>
<td>51.9</td>
<td>0.5</td>
<td>18.8</td>
</tr>
<tr>
<td>1990</td>
<td>MWBM</td>
<td>7.5</td>
<td>0.3</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>NAM</td>
<td>8.5</td>
<td>0.8</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>MMC</td>
<td>35.6</td>
<td>0.1</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Forecasted results above by MWBM and NAM are all based on given historical meteorological data. But in real time flow forecasting, forecasting accuracy is mainly up to two factors, one is the hydrological model accuracy and the other is the meteorological forecasting accuracy. And conclusions obtained above mainly reflect the first factor.
In addition, in flow forecasting with the statistic method, (i.e. the MMC method), long series are mostly needed. However, 40 years is a reasonable long period of data. The forecasted error in this case is mainly due to not using the rainfall data.

Above results denoted that the three models can be used to forecast discharge in the next one or two years. If meteorological data in the future can be forecasted for a longer period and are of high quality, MWBM and NAM should be adopted first.

**Future activities after back China**

Applying Nam model and MWBM in the other basins, and compare simulation results in different basin.

We will do our best to apply the PCRASTER software in China if we can get the digital map of Yellow Rive basin.

We will try to use the Xin’anjiang model for mid-term and long-term flow forecasting in the suitable area of Yellow River basin.

Applying the MMC to other area in Yellow River.
REFERENCES


PIETER H.A.J.M. VAN GELDER, Analysis and statistical forecasting of trends in river discharges under uncertain climate changes

VADIM A. KUZMIN, PIETER H.A.J.M. VAN GELDER, The principles of catastrophic flood forecasting


Papers with Yuanfang Chen

Reference to Benson (1968)


Computational methods in water resources. Vol 1. Modeling surface and sub-surface flows: proceedings of the VII International Conference, MIT, USA, June 1988

Modeling Components of Hydrological cycle: international symposium on rainfall-runoff modeling, Mississippi State, May 1981

Applied modeling in Catchment Hydrology; international symposium on rainfall-runoff modeling, Mississippi State, May 1981


Rainfall-runoff Relationship; international symposium on rainfall-runoff modeling, Mississippi State, May 1981

Hydrological research bains and the environment: proceedings of the International conference, 24-28 September 1990

Model validation: perspectives in hydrological science. Chichester Wiley 2001

Diermanse, Ferdinand lennaert Machiel.: Physically based modeling of rainfall-runoff processes.


Distributed hydrological modeling. Dordrecht Kluwer Academic 1996

Recent advances in the modeling of hydrologic systems. Dordrecht Kluwer 1991
RIENTJES,T.H.M.: Hydrological models; Cthe4431; dictaat behorende bij college
CT4431.Delft TU Delft 1998
Computational methods in water resources IX.Vol. 2. Mathematical modeling in water
resources. Southampton Computational Mechanics Publications 1992
Haasnoot,M.: Impact of climate change and land subsidence on the water systems in the
Netherlands: terrestrial areas . Lelystad RIZA 1999
Parameter identification and inverse problems in hydrology ,geology and ecology.Dordrecht
Kluwer Academic 1996