The Study on Fundamental Properties of Stochastic Variation of Natural Annual Runoff of Main Rivers in China

Shaowen Zhang[1], Wensheng Wang[1], Jing Ding[1], Jie Liao[2], Xing Wan[3]

1. School of Hydraulic Engineering, Sichuan University, Chengdu, Sichuan 610065, China.
2. School of Chemistry, Sichuan Teacher University, Chengdu, Sichuan 610068, China.
3. School of Electron, Chengdu University of Information Technology, Chengdu, Sichuan 610051, China

Abstract: Based on the observed data which were collected from 177 gauging stations, the stochastic properties of sequential variations of natural runoff of main rivers in China has been probed with the theory of stochastic process and technology of time series analysis. The theory and technology can demonstrate many kinds of characteristics of the annul runoff time series; the trend, periodicity, serial correlation, and pure random. The trends of these annual runoff sequences may generally not be significant for large river and the almost periodicities of the sequences are mainly influenced by astronomic and climatic factors. The study also shows that the sequences exist the time dependence and the persistence in many rivers, and the cross-correlation coefficients between two annual runoff sequences vary in a wide range from positive to negative for the four largest rivers in China, but the simultaneous relation of annual runoff between the two largest rivers in China, Yangtse River and Yellow River, appears to be positive. These results are important for feature research of the water resources and environment assessment. [The Journal of American Science. 2005;1(1):46-54].

Key words: main rivers in China; annual runoff; variant trend; variant periodicity; variant dependence and persistence; cross-correlation coefficients

1. Introduction

Sequential variations of annual runoff are very complicated. At the early time of that century they were considered as pure random phenomena and described by using probabilistic distribution curves (Salas et al, 1980). Up to the thirties, a new idea arose that the sequential variations might be implicated in some periodic properties (Chow, 1978). In the beginning of the sixties, the concept that annual runoff might be autocorrelated was introduced into the analysis in the variations of annual runoff (Yevjevich, 1972). Since then, the theory of stochastic process and technology of time series analysis have been applied to developing stochastic models for describing such variations.

The analysis in stochastic properties of sequential variations of annual runoff of main rivers in China has been made in this paper. It is based on the observed data, which were collected from 177 gauging stations. The areas of their catchments range from 10,000 to 200,000 square kilometers. The rivers where these stations are located can be characteristic of representing various regimes of annual runoff variation. The amount of station years for this analysis is about 6100. The length of date is 34.5 years on average. The longest one is 108 years and the shortest is 22 years.

The stochastic properties of annul runoff sequence, such as trend, almost periodicity, serial correlation, and pure random, have been identified and examined by using modern effective methods (Kottegoda, 1980; Sala, 1993).

The final results obtained from the extensive and exhaustive analysis are very useful and important for developments of water resources and environment assessment (Hipel, 1985).

2. Trend of annual runoff sequence

Trend may exist in an annual runoff sequence due to the impact induced by nature or human begins. Several methods and techniques such as Kendall rank correlation testing, Spearman rank correlation testing (Siegel, 1956), regressive coefficient testing, and moving

average analysis (Ding, 1988) have been adopted for investigating whether a trend exist in a given sequence of annual runoff or not. The results from the hypothesis test and analysis have shown that most of the sequences studied do not exhibit any trends. If the sequences with short lengths are not considered, 80 percent of those with sample date more than 40 years do not appear to be any trend. This can be detailed as follows:

(1) For large rivers with long lengths of date, such as Yangtze River and Yellow River, the trends of annual runoff sequences may not be significant.

(2) For some rivers with short lengths of date their annual runoff sequence have shown relatively long-range slow upturns or downturns which are apparent trends (Figure 1), because

they may seem as such for short samples, but in reality they are just part of the sampling variations when are considered for large sample sizes. These apparent overyear trend are not population properties of annual runoff sequence.

(3) For some river with relatively small basin areas, the trends of their annual runoff sequences may be significant due to man-induced activities.

Therefore, in general, the sequences of annual runoff in large basins in China would be considered as non-trend sequences. When the rivers with relatively small basins are concerned, the man-induced activities have to be taken into account, because such activities may perhaps cause a trend in annual runoff sequence.
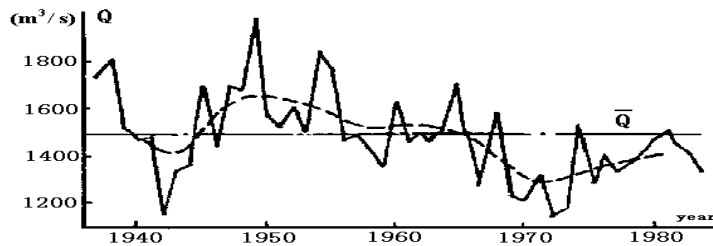

Figure 1. Annual flow sequence at Tong Jiezhi station

3. Almost periodicity of annual runoff sequence

Astronomic cycles are the basic causes of almost periodicities in the characteristics of annual runoff sequences. Such periodicities can be identified by using spectral analysis, maximum entropy spectrum analysis, variance analysis, deviation sum curve method, and moving average method (Ding, 1988, Ding, 1977).

Since various analyses and methods mentioned above are different in their assumptions and applicability, the results from statistical inferences and intuitional judgments are not the same. Therefore, in this study the most significant almost periodicity for a given annual runoff sequence was mainly determined through comprehensive analyses based on the results obtained from various methods and justification in its cause. Additionally, in the

whole analyses the emphasis was put on the rivers with long sample data. The main periodicity obtained in such a way was here referred to as first significant almost periodicity, which will be discussed more detail.

34 stations from all the stations concerned in this study were identified to have significant almost periodicities with periods vary from 3-year to 15-year. These periods for some large rivers in China are listed in Table 1.

It has shown that, in general, the periods are less than 15-year. Most of them are characteristic of short terms, that is, 3-year, 5-year, 7-year, and so on. As far as the periods of longer than 15-year is concerned, it is difficult to inference on them and the results are not reliable because of the limit of observed date.

The almost periodicities mentioned above are characteristic regionalization. The almost periodicities with 3-year period may mostly occur in the upstream region of Yellow River, Wujiang

River, Xijiang River, Yuanjiang River, and so on. Those with 7 to 11 year period may occur in the east region of China. Especially, those with 7 to 9 year period extensively occur in this region. Those with 15-year may occur in some trunk streams of upper reaches in Yangtze River and some of its branches.

The occurrences of the periods stated above may be caused by natural factors. The 3-years period, which has occurred in the west part of China, is caused by climatic factors, which have significant 3-years period. This can be illustrated by Figure 2, in which two sample sequences were plotted.

Sequence (a) is annual precipitation at Lanzhou; Sequence (b) is annual runoff at Shan County Station on Yellow River. Figure 2 indicates the variation of annual precipitation may lead to the variation of annual runoff (Xu, 1982).

The 7-15 year periods may be the comprehensive results from the 11-year period of the solar spot number and winter temperature at

north Pacific Ocean, the 6-11 year periods of frequency number of typhoon occurrence and other short periods of some factors such as the index of atmospheric circulation.

In summary, the sequences of most river in China do not exhibit any periodicities, only small number of them appear to be some periodicities, which are characteristic of regionalization. The variances explained by the almost periodic variations, in general, are not large. Thus, the variations of annual runoff sequences of main rivers in China are less affected by periodic factors.

4. Time dependence of annual runoff sequence

The annual runoff in a river may be autocorrelated. This property is referred to as the time dependence of annual runoff sequence. It is caused mainly by storage capacity of that river (Klemes, 1978). In general, an autocorrelation coefficient can be used as the index to measure the degree of such dependence.

Table 1. Periods of Annual Sequences in Some Stations

| River name	Station name	Basin Area (sq . kilo.)	Sample size (year)	Period (year)
Yellow	Shan County	687,000	66	3
Yellow	Lanzhou	223,000	50	3
Datong	Xangtang	15,100	45	3
Yaohe	Lija	19,700	38	3
Weihe	Hua County	10,600	50	3
Jinghe	Zhangju	43,200	53	3
Wujiang	Wulong	83,200	34	3
Hongshui	Qianjiang	128,000	33	3
Lioujiang	Liouzhou	45,700	39	3
Luanhe	Luan County	43,100	51	5
Chaobai	Souzhuang	17,600	62	5
Baijiang	Sijiao	38,300	33	7
Xiang	Xiangtan	81,600	35	7
Ganjiang	Waizhou	80,900	35	7
Xinjiang	Huzhou	161,400	35	7
Qujiang	Yuren	13,600	35	8
Huaihe	Mongchen	15,500	35	8
Hanjiang	Huan	29,000	34	8
Qujiang	Loduxi	38,100	32	9
Hanshui	Huangjiagang	95,200	56	9
Shahe	Lohe	12,600	35	11
Yinghe	Fuyang	38,200	35	11
Yangtze	Cuntan	867,000	91	15
Yangtze	Yichang	1,000,500	108	15
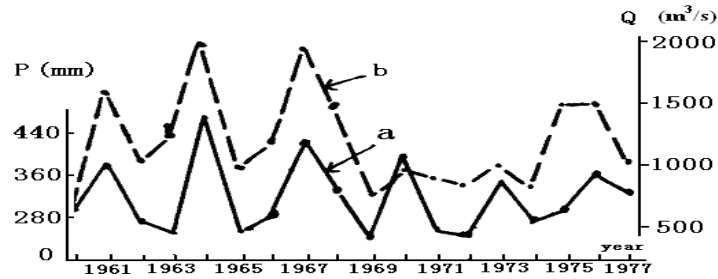
Figure 2. The relation of annual streamflow at Shan County station with annual precipitation Lanzhou station

(1) General properties of time dependence

The autocorrelation coefficient $r(k)$ of the order k, or lag k, is defined by the product-moment correlation coefficient between the member of a given sequence that are k items, or k lags, or k intervals apart. The autocorrelation coefficients of order one, $r(1)$, computed from the data at 177 stations show a great variation range from -0.461 to 0.535. The values of $r(1)$, at 119 stations are positive and those at 58 stations are negative. The $r(1)$ can be ranked into four types:

Type A, $r(1) \geq 0.3$, represents the strong dependence of annual runoff sequence.

Type B, $0.1 \leq r(1) < 0.3$, represents the weak dependence of annual runoff sequence. However such dependence has significant impact on planning and designing of water resources engineering (Qing, 1984).

Type C, $0 \leq r(1) < 0.1$, represents the minor dependence of annual runoff sequence.

Type D, $r(1) < 0$, represents the negative dependence of annual runoff sequence.

Some statistical information of these four types is in Table 2.

Table 2. Types of r(1)

Type	Range of r(1)	Number of station	Percentage of all station (%)	Average value of r(1)
A	≥ 0.3	38	21.5	0.34
B	$0.1 \sim 0.3$	45	25.4	0.19
C	$0.0 \sim 0.1$	36	20.3	0.05
D	< 0	58	32.8	-0.14

Table 2 indicates that the time dependence of annual runoff sequences do exist in almost half of all the sequences involved in this study.

The correlograms of various annual runoff sequences are different. For the four types mentioned above, the properties of their corresponding correlograms are listed in Figure 3 (four stations were selected as representatives for

the four types). The correlograms of Tape A and Tape B will decay with the increase of k, but their decaying speeds are different, the former is slower than the latter. The correlograms of Type C and Type D will fluctuate around the horizontal axis, but the former varies in a smaller range than the latter.
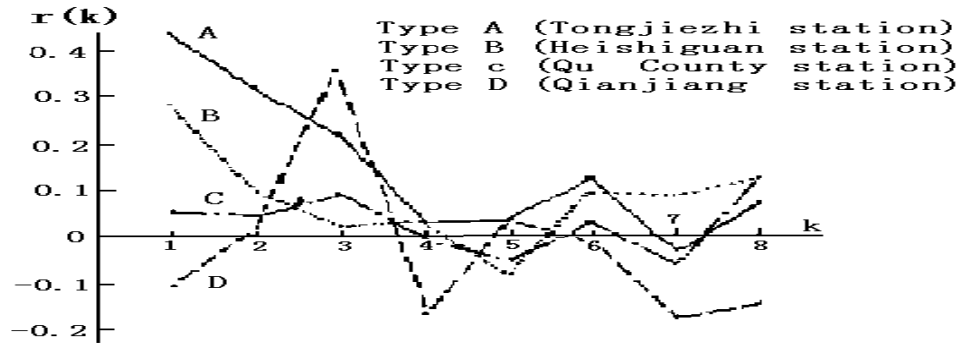

Figure 3. Correlograms of annual runoff sequence for the four representative stations

(2) The regional properties of time dependence

The autocorrealtion coefficients of order one, r(1), appear to have regional properties. The whole country can be roughly divided into three subareas according to the values of r(1). These subareas are: the relatively large value subareas for r(1)≥0.2, the small value subareas for 0.0 ≤ r(1) ≤ 0.2 , and the negative value subareas for r(1)<0 (see Appendix Figure 1). The configuration of the subareas in this figure may be jointly made by comprehensive effect of storage capacities of basins and periodic factors. When the periodicity in a sequence is not significant or the influence of periodic climatic factors on r(1) is weak, the greater the storage capacity is, the large value of r(1) will be.

Therefore, the large value subareas may be located in the region where basins are characteristic of great storage capacity, while the small value subarea may be located in the region where basins are characteristic of small storage capacity. As far as negative value subarea is concerned, although the value of r(1) is also influenced by the storage capacity, it may be mainly influenced by 3-years significant period of annual runoff sequence which will further be induce by same period of annual precipitation sequence. Figure 4 illustrates this situation.

The correlograms for the three subareas are different. Figure 5 indicates such difference, where r(1) represents the average computed from all station in a corresponding subarea.
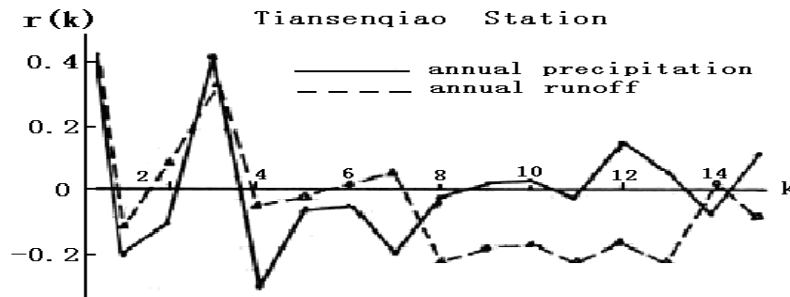

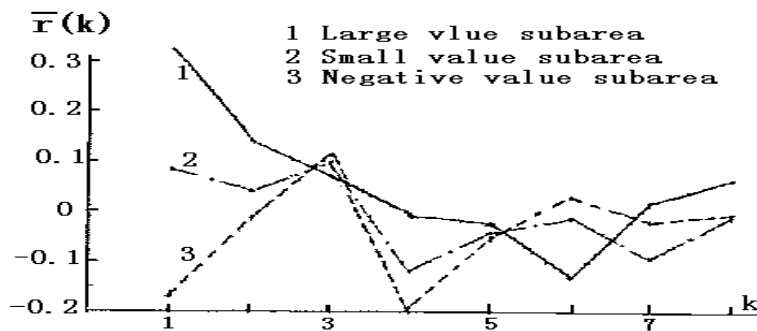**Figure 4. Correloggram of annual runoff and precipitation sequence**

Figure 5. Correlograms for the three subbarea

5. The persistence of annual runoff sequence

The important characteristic of annual sequence is the persistence. This is the property by which high flows tend to follow high and low flows tend to follow low flows. Several methods and techniques have been put forward to

investigate such persistence. Run analysis, Hurst coefficient analysis, transition probability method, and drought parameter method were chosen in this study.

In run analysis the mean value of a given sequence was taken as a crossing level and the emphasis of investigation was put on mean negative run length (EL) which can represent the

persistence characteristic of draught. For Hurst coefficient k, its estimates were obtained by a graphical procedure proposed by Mandelbrot and Wallis (1969). The drought parameter was proposed by D. M. Hershfield (1977). The following equation can be used to estimate the probability:

$$P= q^{j-1}(1-q) \qquad (1)$$

Where p is the probability by which the drought with j consecutive years occurs. The q is defined as drought parameter representing probability by which the drought occurs on the condition one drought has occurred in the preceding year. This parameter can be estimated by the equation.

$$q= 1- \frac{S}{D} \qquad (2)$$

Where S denotes total number of negative runs in a given sequence; D denotes the sum of all negative run lengths in a given sequence.

In order to estimate transition probability, three states of annual runoff were defined as follows:

Q>Q30% for high annual runoff;

Q30%>Q>Q70% for normal annual runoff;

Q<Q70% for low annual runoff.

Where Q denotes annual flow; Q30% and Q70% stand for annual flows with frequency of 30% and 70% respectively.

The results from the comprehensive analysis and computation based on observed data at 177 stations shows:

(1) Hurst coefficient k varies from 0.450 to 0.877, with a mean of 0.72, which is almost equal to the value of 0.73 obtained by Hurst from 800 sequences (Kottegoda, 1980).

(2) Average negative run length EL varies from 1.3 to 4.3, with a mean of 2.39.

(3) Drought parameter q varies from 0.250 to 0.736, with a mean of 0.546.

(4) k, EL, and q are, to a considerable degree, related to r(1).

It has been pointed out that r(1) has regional properties, therefore, the persistence parameters of k, EL, and q exhibit the same property as the r(1) does. They have been shown in Appendix Figures 2, 3, and 4. Moreover, Table 3 indicates the average values of these parameters for the three subareas.

Annual runoff in China tends to decrease from southeast to northwest and, consequently, there exist three zones, namely, humid zone, transitional zone, and arid zone.

The k, EL, and q associated with the three zones are listed in Table 4, which indicated that the persistence parameters of k, El, and q tend to decrease with the increase of annual runoff, which consists with the climatic properties in China.

6. The dependence between annual runoff sequences

The dependence between two runoff sequences can be determined by computing the cross-correlation coefficient expressed by $r_{xy}(k)$, between the two sequences x, and y (Yevjevich, 1972). The coefficients between the four largest rivers in China are listed in Table 5. From the table we can see:

(1) The simultaneous relation of annual runoff between Yangtze River and Yellow river appears to be positive .

(2) The relation between Songhua River and Xijiang River appears to be negative.

Table 3. Average values of persistence parameters for the three subarea

Subareas	r(1)	k	EL	Q	P_{SS}(%)	P_{EE}(%)	P_{LL}(%)
Large value	0.322	0.792	2.87	0.625	45.8	41.2	48.9
Small value	0.085	0.703	2.34	0.556	34.7	37.8	33.0
Negative value	-0.170	0.635	1.87	0.448	22.0	40.4	26.6

Note: P_{SS}, P_{EE} and P_{LL} stand for the transition probabilities by which low, normal, and high annual runoffs will transit to the same state respectively.

Table 4. k, EL, and q for the three zones

Zones	K	EL	q
Humid	0.695	2.04	0.51
Transition	0.736	2.65	0.586
Arid	0.800	3.01	0.656

Table 5. Cross-correlation coefficients

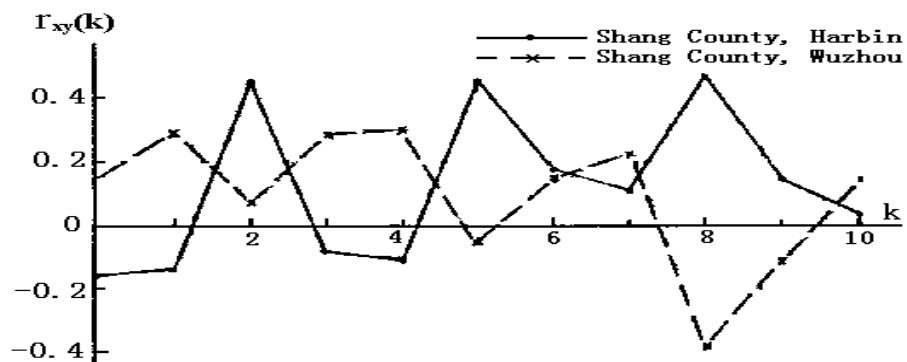| River	Lag 1	Lag 2	Lag3	Lag4
Yangtze —Yellow	0.286	0.213	0.06	-0.013
Yangtze —Songhua	0.023	0.184	0.203	0.107
Yangtze —Xijiang	0.01	0.062	0.225	0.008
Yellow —Songhua	-0.158	-0.144	0.451	-0.086
Yellow —Xijiang	0.141	0.293	0.073	0.292
Songhua —Xijiang	-0.346	-0.161	-0.166	-0.392

Figure 6. Cross-correlogram of annual runoff sequences at Shan County station (Yellow River), Harbin station (Songhua River), and Wuzhou station (Xijiang River)

The cross-correlogram between Yellow River and Songhua River and that between Yellow River and Xijiang River are plotted in Figure 6, which shows that the variation of $r_{xy}(k)$ with lag k is very complicated.

7. Conclusions

Final conclusions would be drawn as follows:

(1) For large river, the trends of their annual runoff sequences may generally not be significant.

(2) The almost periodicities of annual runoff sequences are mainly induced by astronomic and climatic factors. These periodicities appear only for a small number of the rivers concerned in this study, and their periods have been indicated to be not more 15 years.

(3) For almost one half of all the rivers involved in this study, the time dependence of their annual runoff sequences exists. The values of r(1) mainly representing such dependence can be ranked into four types. Moreover, r(1) appears to have regional properties.

(4) The persistence of annual runoff sequence exists in many rivers. The persistence characteristic can be described by average negative run length, Hurst coefficient, drought parameter, and transition probability. These persistence parameters exhibit the same regional properties as autocorrelation coefficient of order one does.

(5) For the four largest rivers in China, the

cross-correlation coefficients between two annual runoff sequences vary in a wide range, from positive to negative. The simultaneous relation of annual runoff between the two largest river in China, Yangtse River and Yellow River, appear to be positive.

Correspondence to:

Shaowen Zhang
School of Hydraulic Engineering
Sichuan University
Chengdu, Sichuan 610065, China.
E-mail: Zsw.scu2002@163.com or
Zsw.scu2002@126.com

References

[1] Chow VI. Evolution of Stochastic Hydrology, In Application of Kalmam Filter to Hydrology, hydraulics and Water Resources, ed. Chao-lin, Chiu, University of Pittsburgh, 1978:60─100.

[2] Ding J, Deng Y. Stochastic Hydrology, Press of Chengdu University of Science and Technology, 1988:66─73, in Chinese.

[3] Ding J, Liu QS. Stochastic Hydrology, Press of Water Resources and Electric Power in China, 1977:58─105, in Chinese.

[4] Hershfield DM, et al. Some Measures of agricultural Drought. Floods and Droughts.Water Resources (Publication Colorado), 1972:50─101.

[5] Hipel KW. Time Series Analysis in perspective. Water Resources (Bulletin) 1985;21(4):95─104.

[6] Klemes V. Physically Based Stochastic Hydrologic Analysis, Advances in Hydrology 1978; 11:97─103Kottegoda NT. Stochastic Water Resources Technology(John Wiley and Sons, New York), 1980:10─51.

[7] Mandelbrot BB, Wallis JR. Some Long-run Properties of geophysical records, Water Resource Research 1969;5:201─209.

[8] Qin AL. The Preliminary Research on Autocorrelation Coefficients of Annual Runoff Sequence of Main Rivers in China, Hydrology, 1984;6:56─72, in Chinese.

[9] Salas JD, et al. Applied Modeling of Hydrologic Time Series, Water Resources (Publication, Colorado) 1980:12─16.

[10] Salas JD. Analysis and modeling of hydrologic time series, in Hand book of Hydrology, edited by D. R. Maidment, Chapter 19, Mcgraw-Hill, New York, 1993.

[11] Siegel S. Nonparameteric Statistics for the Behavioral Sciences (Mcgraw-hill Book Company, Inc.), 1956:89─92.

[12] Xu G, Dong A. Three-year Almost period of Western Precipitation in China, Metrology in Tableland 1982;1(2):81─90, in Chinese.

[13] Yevjevich, V., Stochastic Process in Hydrology, Water Resources Publications, Colorado, 1972.

Appendix Figure 1
Regionalization of first order autocorrelation coefficient r(1)

Appendix Figure 2

Regionalization of
Hurst Coefficient k

Songhua

0. 809   River

0. 825

0. 676

0. 802

0. 615   River

Yellow

0. 757   0. 703

River

Yangtze

0. 771   Xijiang   0. 639

River

Appendix Figure 3

Regionalization of average
Negative run lengh EL

Songhua

2. 42   River

3. 23

2. 48

3. 26

1. 83   River

Yellow

2. 54   2. 42

River

Yangtze

Xijiang   1. 88

2. 01   River

Appendix Figure 4

Regionalization of
drought parameter q

Songhua

0. 544   River

0. 673

0. 578

0. 447   0. 677

River

Yellow

0. 59   0. 527

River

Yangtze

0. 497   Xijiang   0. 448

River