Statistical Analysis of Long-term Series Rainfall Data: A Regional Study in Southeast Asia

J.P. Bricquet[1], A. Boonsaner[2], B. Bouahom[3] and T.D. Toan[4]

Abstract

The occurrence of many extreme events in hydrology cannot be forecast on the basis of deterministic information without sufficient skill and lead time. In such cases, a probability approach is required in order to incorporate the effects of such phenomena into decisions.

A short statistical study was conducted on a long-term series of data on daily rainfall from three weather stations in Vietnam (Hoa Binh), Laos (Luang Prabang), and Thailand (Phrae). These stations are relatively close to the MSEC study sites. The recorded period ranges from 28 to 50 years. The analysis was made on annual and maximum daily rainfall. Statistical adjustments gave the values for decennial and centennial return periods. Annual rainfall followed a normal distribution while maximum daily rainfall followed a Pearson 3 distribution.

All stations had reached annual values close to the millennium frequency (frequency 0.99) for the maximum values recorded and close to the centennial frequency (0.01) for the minimum values. All stations have reached the maximum daily values equal or higher than the centennial value. The highest observed value for Hoa Binh station (416.4 mm) is probably a millennium occurrence. There was a small tendency of the rainfall values to decrease in 1997 which was a strong El Niño year in the region. From this study, we may adjust rainfall data inputs for the erosion model developed by MSEC and extend the study to a larger area.

Introduction

The occurrence of many extreme events in hydrology cannot be forecast on the basis of deterministic information without sufficient skill and lead time. In such cases, a probability approach is required to incorporate the effects of such phenomena into decisions. If the occurrence can be assumed to be independent of time, then frequency analysis can be used to describe the likelihood of any one or combination of events over the time horizon of a decision (WMO, 1983).

Interpretation of precipitation has two major purposes. One purpose is to evaluate the observations that sample a precipitation event or series of events. The evaluation of the observed sample includes consideration of extraneous influences, such as deficient or changing

[1] International Water Management Institute, Southeast Asia Regional Office, Bangkok, Thailand
[2] National Parks, Wildlife and Plant Conservation Department, Bangkok, Thailand
[3] National Agriculture and Forestry Research Institute, Vientiane, Lao PDR
[4] National Institute for Soils and Fertilizers, Hanoi, Vietnam

gauge exposure, and interpretation of the effects of physical environment, such as physiography. The other purpose is to describe the event in a form appropriate for display, subsequent analysis, or other applications.

Little is known about trends in rainfall in the Southeast Asia region. A study conducted by Manton *et al.* (2001) showed a gap in the Indochinese Peninsula. Via the Management of Soil Erosion Consortium (MSEC) led by the International Water Management Institute (IWMI), it has been possible to gain access to three recorded series in this region. This paper provides a statistical analysis of long-term rainfall data in the region.

Location of the Study

The analysis was made using the data from three weather stations relatively close to the MSEC project sites. These stations are Hoa Binh in Vietnam (20° 49' N, 105° 20' E), Luang Prabang in Lao PDR (19° 53' N, 102° 08' E), and Phrae in northern Thailand (18° 08'N, 100° 10'E) (Figure 1). The period of record is 41 years for Hoa Binh, 51 years for Luang Prabang, and 28 years for Phrae. Daily rainfall data are available in all stations.

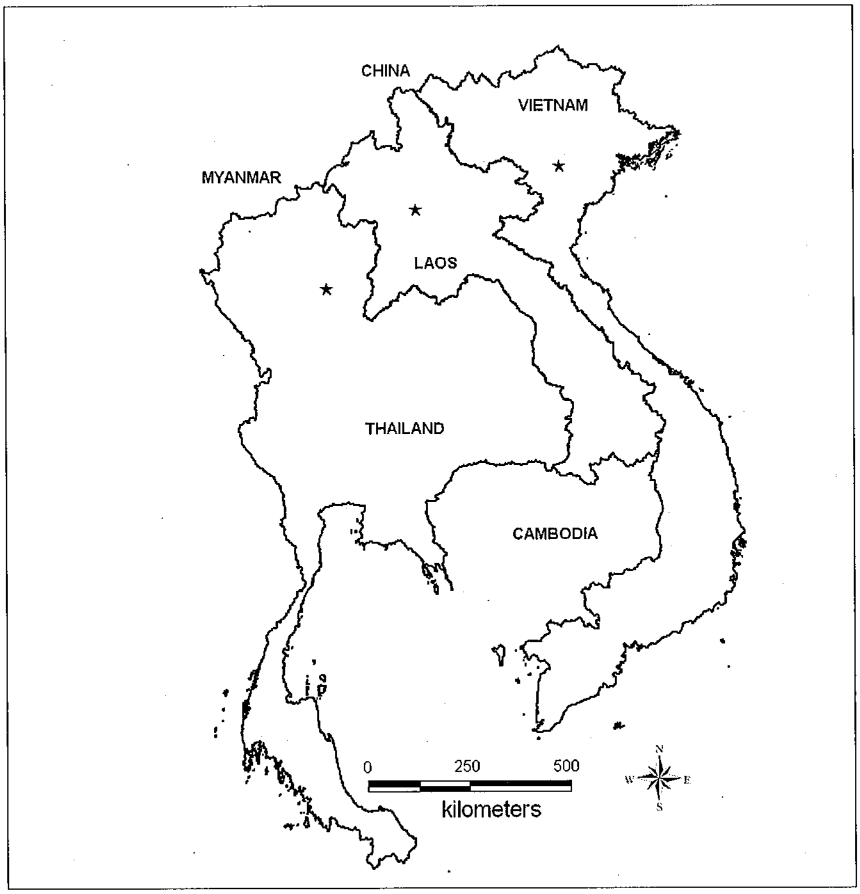

Figure 1. Location of the three stations

Results

Homogeneity of the Data

The homogeneity of hydrological or meteorological data is a requirement for valid statistical application. The most commonly used information about non-climatic influences comes from records of station movement, changes in instrumentation, problems with instrumentation, sensor calibration, changes in surrounding environmental characteristics, observation practices, and other similar occurrences (Guttman, 1998). The double-mass curve analysis introduced by Kohler (1949) is a graphical method of identifying and adjusting inconsistencies in a station record by comparing its time trend with those of other stations. Changes in slope of the double-mass curve may be caused by changes in exposure or location of gauge, change in procedure in collecting and processing data, etc (WMO, 1994). The data collected from all the sites within the region should be highly correlated, have similar variability, and differ only by scaling factors and random sampling variability.

As shown in Figure 2, there is no change in the slope of the curve for both the Hoa Binh and Phrae stations. So, we can consider all series as homogenous. With the very close correlation coefficient, it is also possible to extend the series of Hoa Binh and Phrae but only for annual or monthly values.

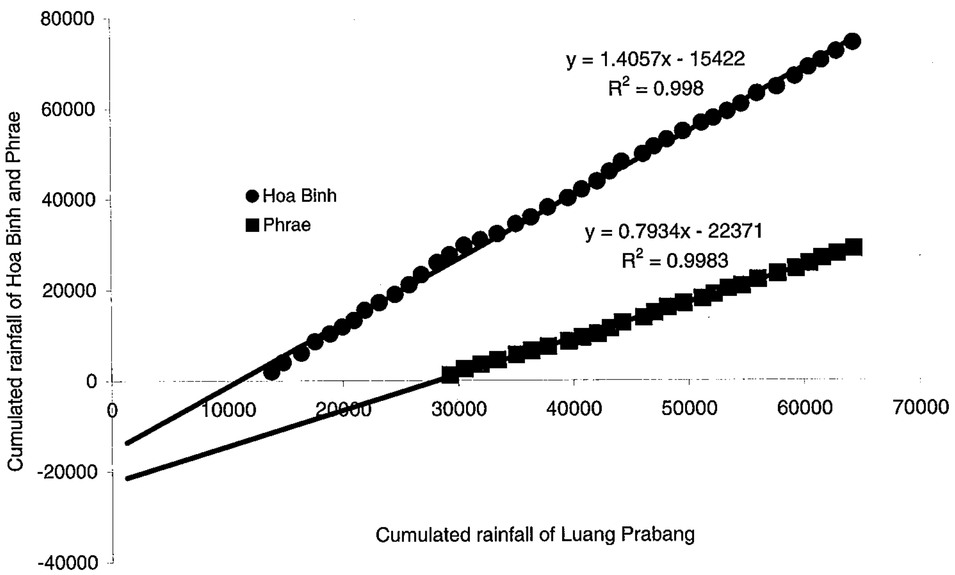

Figure 2. Double-mass curve for Hoa Binh and Phrae stations

Annual Rainfall Analysis

"Climatic Normal" is defined by the World Meteorological Organization (1983) as "period averages of a climatic element such as temperature or precipitation computed for a uniform and relatively long period comprising at least three consecutive ten-year periods". Manton *et* *al.* (2001) assumed that the annual total rainfall had generally decreased between 1961 and 1998. And the number of rainy days (with at least 2 mm of rain) has decreased significantly in Southeast Asia. Looking at the variations of the annual total rainfall (Figure 3) of the three stations, it is difficult to detect a trend of an increase or decrease in the annual amount of rain. Possibly a small decrease can be detected but not quantified now.

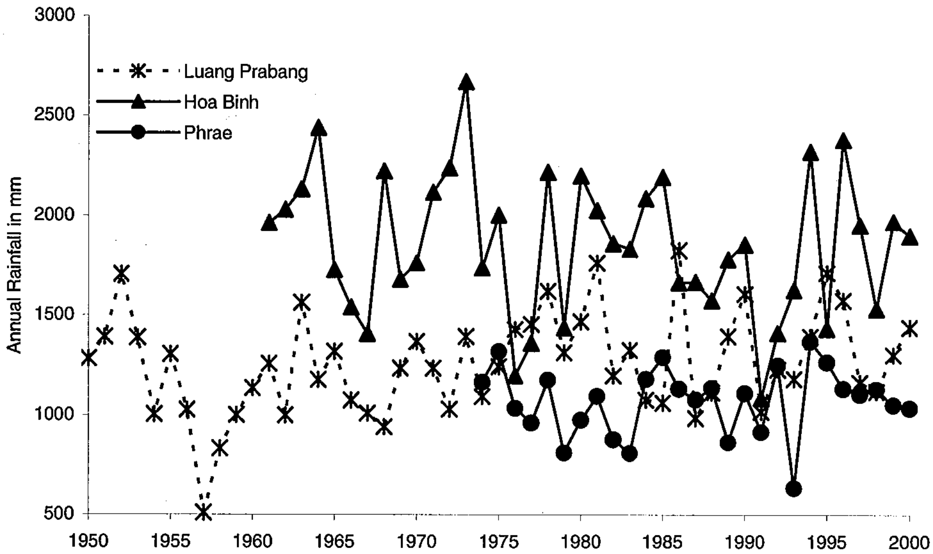

Figure 3. Variation of the annual rainfall at the three stations

The 1957 value of Luang Prabang looks very low. Checking the record, we cannot reject this very small amount of rain (511.1 mm). We can assume the same comment for the 1993 value for Phrae (635.9 mm). Standard statistical results are presented in Table 1.

Table 1. Summary statistics of the annual rainfall of the three stations

	Average	Variance	Std. deviation	Min.	Max.	Std. kurtosis
Hoa Binh	1,856.7	127, 290	356.8	1,085.2	2,671.7	-0.473
Luang Pranbang	1,261.2	65, 058	255.1	511.1	1,827.5	0.817
Phrae	1,084.6	34, 107	184.7	635.9	1,461.3	0.273

Annual rainfall usually follows a Gauss statistical distribution (Figure 4a, 4b). Figure 4a shows the statistical adjustment, while Figure 4b presents the frequency distribution. Accordingly, we calculated the different return period values with the Normal law. Results are presented in Table 2.

According to the calculations, we can see that all stations had already reached annual values close to the millennium frequency (frequency 0.99) for the maximum values recorded. For the minimum values, all stations have already reached centennial frequency (0.01).

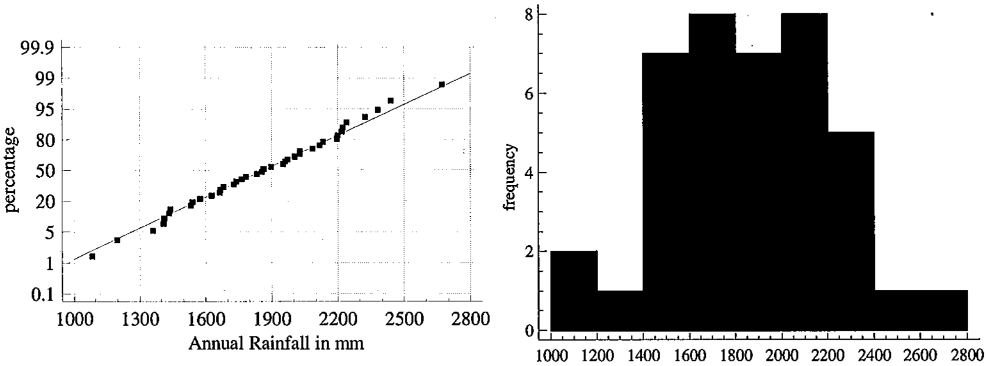

Figure 4. Annual rainfall at Hoa Binh station: statistical adjustment (left), frequency distribution (right)

Table 2. Frequency analysis of the annual rainfall of the three stations

	0.01	0.1	0.5	0.9	0.99
Hoa Binh	1,064.2	1,420.8	1,856.7	2,292.7	2,649.3
Luang Pranbang	710.7	958.4	1,261.2	1,563.9	1,811.6
Phrae	654.9	848.3	1,084.6	1,321.0	1,514.3

Maximum Daily Rainfall

Figure 5 presents the distribution of the maximum daily rainfall for the three stations. For this type of data, we have no Gauss distribution and the data can be adjusted with a Log Normal distribution (Figure 6). Standard statistics are given in Table 3.

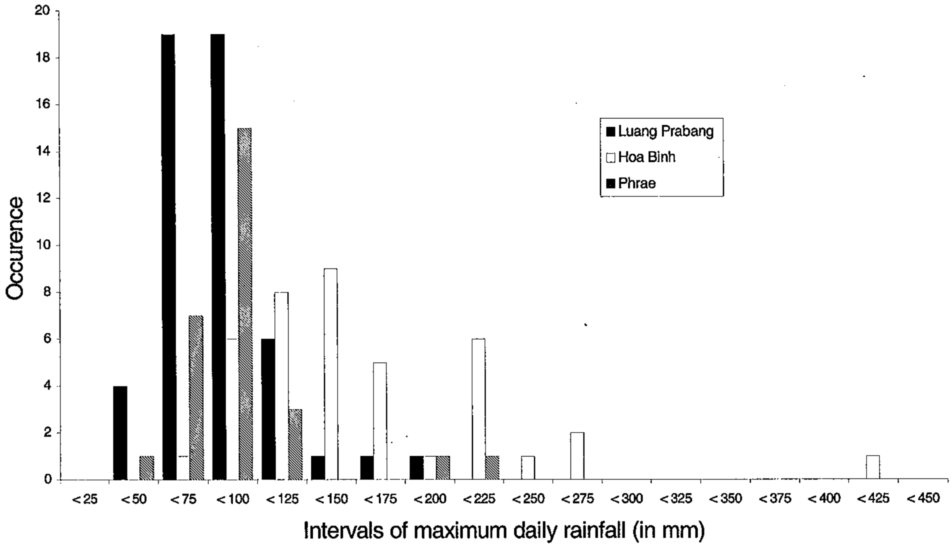

Figure 5. Frequency histogram of maximum daily rainfall at the three stations

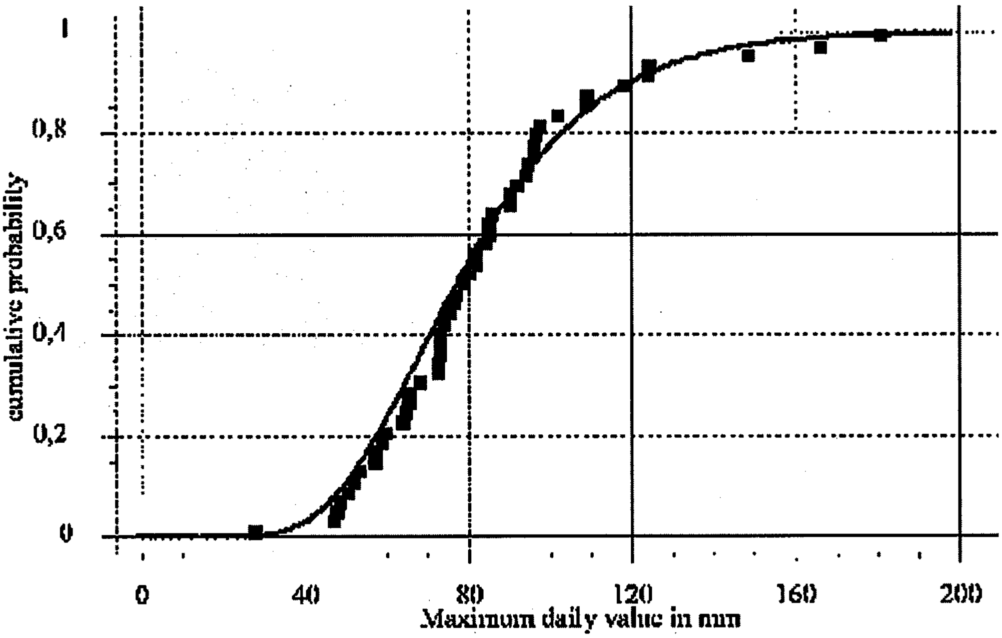

Figure 6. Adjustment of maximum daily rainfall at Luang Prabang station

Table 3. Summary statistics of the maximum daily rainfall of the three stations

	Average	**Variance**	**Std. dDeviation**	**Min.**	**Max.**	**Std. kurtosis**
Hoa Binh	156.1	4, 293.2	65.5	58.3	416.4	6.620
Luang Pranbang	83.2	852.8	29.2	27.7	180.7	3.615
Phrae	90.2	1, 278.8	35.7	43.3	218.2	7.062

The calculations of return period values are done following the Pearson 3 law. Results are given in Table 4. Pearson 3 law is one of the statistical laws used in hydro-meteorological studies especially for unique events.

Table 4. Frequency analysis of the maximum daily rainfall of the three stations

	0.01	0.1	0.5	0.9	0.99
Hoa Binh	64.4	90.2	142.6	238.6	378.9
Luang Pranbang	34.7	50.7	79.2	121.0	168.1
Phrae	48.4	58.0	81.3	132.4	222.7

The highest observed value for Hoa Binh station (416.4 mm) is probably a millennium occurrence. We can also notice that all stations have reached maximum daily values equal to or higher than the centennial calculation. The same comment can be made for the lower values. We could see here an increase in heavy precipitation events as have occurred in the United States (Karl and Knight, 1998).

Conclusion

This study is just a starting point for a deeper investigation of the climate in the region. As we can detect a decrease in the annual total rainfall for the three stations, we should continue this study with more stations in the region, including South China. A limitation of such a study is the low spatial density of stations with homogenous data

Acknowledgement

This study is a contribution to the Management of Soil Erosion Consortium led by the International Water Management Institute and funded by the Asian Development Bank (RETA 5803). Special thanks go to the National Departments of Meteorology of the different countries which provided the data.

References

Guttman, N.B. 1998. Homogeneity, data adjustments and climatic normals. 7th International Meeting on Statistical Climatology, Whistler, Canada, May 25-29, 1998. In Abstract Booklet.

Karl, T.R.; and Knight, R.W. 1998. Secular trends of precipitation amount, frequency, and intensity in the United States, *Bull. Amer. Meteor. Soc.*, 79, 1107-1119.

Kohler, M.A. 1949. Double-mass analysis for testing the consistency of records for making adjustments. *Bull. Amer. Meteor. Soc.*, 30, 188-189.

Manton, M.J.; Della-Marta, P.M.; Haylock, M.R.; Hennessy, K.J.; Nicholls, N.; Chambers, L.E.; Collins, D.A.; Daw, G.; Finet, A.; Gunawan, D.; Inape, K.; Isobe, H.; Kestin, T.S.; Lefale, P.; Leyu, C.H.; Lin, T.; Maitrepierre, L.; Outprasitwong, N.; Page, C.M.; Pahalad, J.; Plummer, N.; Salinger, M.J.; Suppiah, R.; Tran, V.L.; Trewin, B.; Tibig, I. and Yee, D. 2001. Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific, *Int. J. Climatol.* 21, 269-284.

World Meteorological Organization. 1994. Guide to hydrological practices, WMO-No 168.

World Meteorological Organization. 1989. Calculation of monthly and annual 30-year standard normal. WMO-TD/No. 341. Geneva. 11 p.

World Meteorological Organization. 1983. Guide to climatological practices, WMO-No 100.