Chapter 3

Preliminary Results

In this chapter, the results of the preliminary analysis are presented. These results may be classified into three sections as follows.

1. Histograms displaying the data distributions and numerical summaries of all the variables measured in this study.

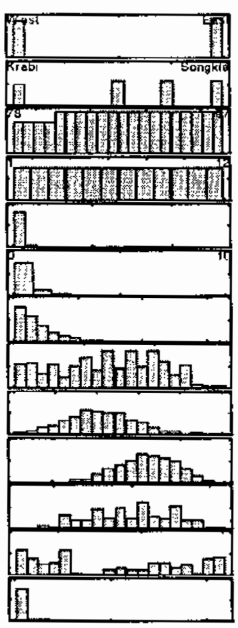
2. Comparison of the outcome (incidence of DHF) and determinant (rainfall) by location (the west and east coasts in the southern part of Thailand).

3. The association between rainfall and DHF incidence, using scatter plots and bivariate time series analysis.

Distributions of Variables

Histograms and extended numerical summaries of each variable are shown in Figure 3.1.

Figure 3.1: Histograms and numerical summaries of data



col	variable	size	mean	st dev	SE	skew	kurt	min	max
1	coast	912	0.526	0.5	0.017	-0.103	1.007	0	1
2	province	912	67.368	1.423	0.047	-0.556	2.086	65	69
3	Year	912	87.921	5.606	0.186	-0.064	1.859	78	97
4	Month	912	6.5	3.454	0.114	0	1.783	1	12
5	cases	912	72.289	183.484	6.076	6.887	70.384	0	2653
6	deaths	912	0.34	0.978	0.032	4.859	33.57	0	10
7	rainfall(cm)	912	17.717	17.486	0.579	2.032	10.404	0	164.05
8	rain days	912	13.331	7.266	0.241	-0.16	2.023	0	30
9	max temp	912	321.397	16.563	0.548	0.186	2.896	280	381
10	min temp	912	236.135	11.758	0.389	-0.399	3.095	186	267
11	humidity	912	79.755	5.379	0.178	-0.138	2.178	65	92
12	pop(1000s)	912	829.224	435.586	14.424	0.144	1.476	239	1512
13	inc/1000	912	0.093	0.225	0.007	5.212	35.826	0	2.203

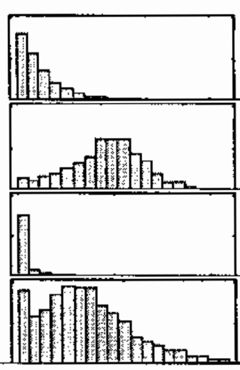
Study of DHF in S Thailand 1978-97

Looking at the histogram for the variable 'year', we see that the data are incomplete for the first four years. This is because no data were available for Krabi until 1982. We also see that the histograms for the monthly numbers of cases and deaths are skewed to the right, and the corresponding skewness coefficients are 6.9 and 4.9, respectively. Since the populations of the four provinces vary substantially (from 239,000 for Krabi to 1,512,000 for Nakhon Si Thammarat), the case incidence rate is a more useful measure of outcome, but this distribution is also substantially skewed, with skewness coefficient of 5.2.

The distribution of rainfall is also positively skewed, with skewness coefficient of 2.0. However, the distributions of maximum and minimum temperature (measured in tenths of degrees Celsius), number of days of rain in the month, and relative humidity, are all approximately symmetric. The minimum temperature recorded over the period was 18.6°C, while the maximum was 38.1°C. The relative humidity varied from a minimum of 65 to a maximum of 92.

Taking appropriate transformations of the data can reduce the skewness in the distributions. Figure 3.2 shows these transformed distributions for rainfall and case incidence/1000. The rainfall is transformed by taking cube roots of the amounts in hundredths of centimeters, and the incidence is transformed by taking logarithms to base 10. To avoid losing data where the number of cases is zero, 0.005 is added to each incidence rate before taking logarithms.

Figure 3.2: Histograms and numerical summaries of transformed data



col	variable	size	mean	st dev	SE	skew	kurt	min	max
1	rainfall(cm)	912	17.717	17.486	0.579	2.032	10.404	0	164.05
2	cubrt(rf)	912	10.61	4.345	0.144	-0.282	2.985	0	25.409
3	inc/1000	912	0.093	0.225	0.007	5.212	35.826	0	2.203
4	log(inc)	912	-1.455	0.564	0.019	0.635	3.117	-2.301	0.344

Study of DHF in S Thailand 1978-97

As a result of the transformations, the skewness coefficient in the rainfall distribution is reduced to –0.282. The corresponding kurtosis is also reduced from 10.4 to 3.0. Since the kurtosis coefficient of a normal distribution is 3, this means that the distribution of the cube roots of the monthly rainfall is approximately normal.

However, the transformation is less effective in removing the skewness of the case incidence distribution. The result of the transformation is to give a skewness coefficient of 0.63 and a kurtosis of 3.17. Clearly, the distribution of the log-transformed incidence has a 'lump' corresponding to zero incidence.

Comparisons of DHF incidence and rainfall by location

Figure 3.3 shows the trend of incidence of DHF in each year between the west coast and the east coast, for each pair of provinces in the southern Thailand (after transforming the monthly incidence rates as described in the preceding section). Similarly, Figure 3.4 shows the trend of incidence of DHF in each year between the northern and southern locations, for the east and west coasts.

These graphs show that the overall incidence of DHF increased during the first decade, and then stabilized. In the southern provinces, there is some evidence that the incidence rate was higher on the east coast, but there is no apparent difference between the east and west coast northern provinces.

From Figure 3.3, in the south location during 1978-1984 the east coast had a greater incidence of DHF than the west coast. And during the period 1985-1997 the incidence of DHF of the east and the west coasts is not different. In the north location, the incidence of DHF between the west and east coasts are not different. But in the year 1991 the incidence of DHF between the west and east coasts is clearly different, possibly due to heavy flooding in this year on the east coast.

Figure 3.3: The incidence of DHF by year and location (east and west) during 1978 – 1997

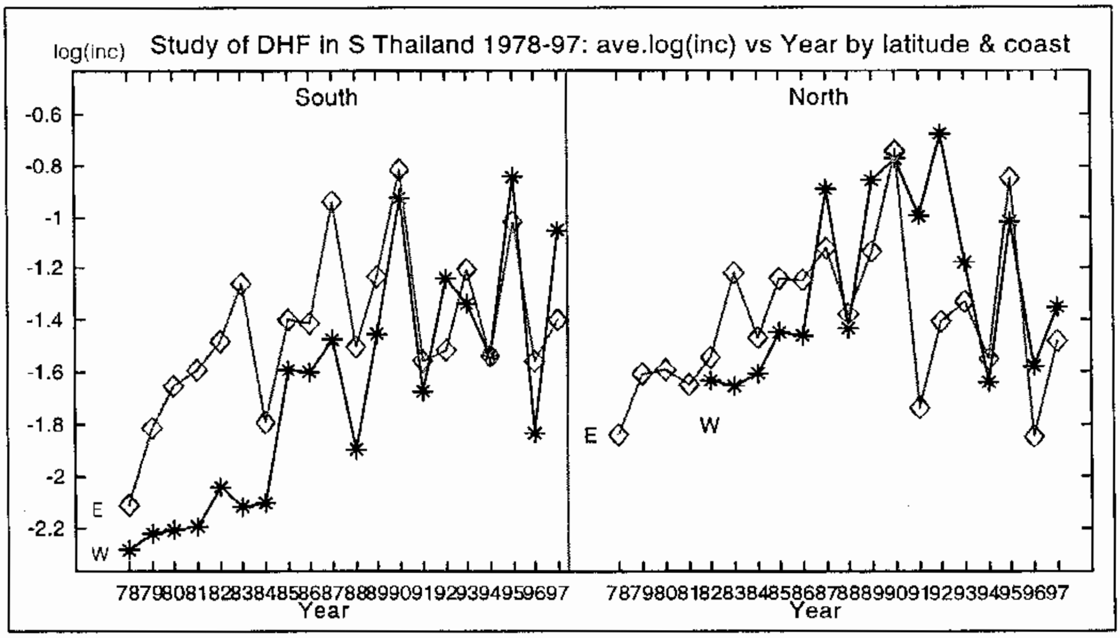


Figure 3.4: The incidence of DHF by year and location (north and south) during 1978 - 1997

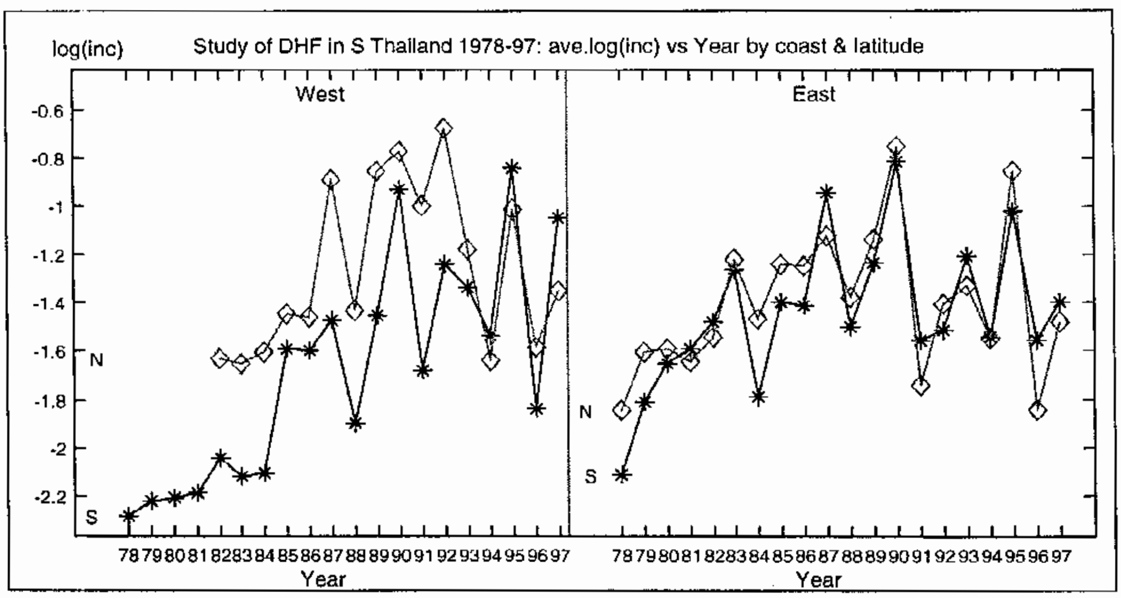


Figure 3.5 presents the seasonal trend of incidence of DHF in each month between the west and east coasts, for each pair of province in the Southern Thailand (again after transforming the data). Similarly, Figure 3.6 presents the seasonal trend of

incidence of DHF in each month between northern and southern location, for the east and west coasts.

Figure 3.5: The incidence of DHF by month during 1978 – 1997 for the east and west coasts

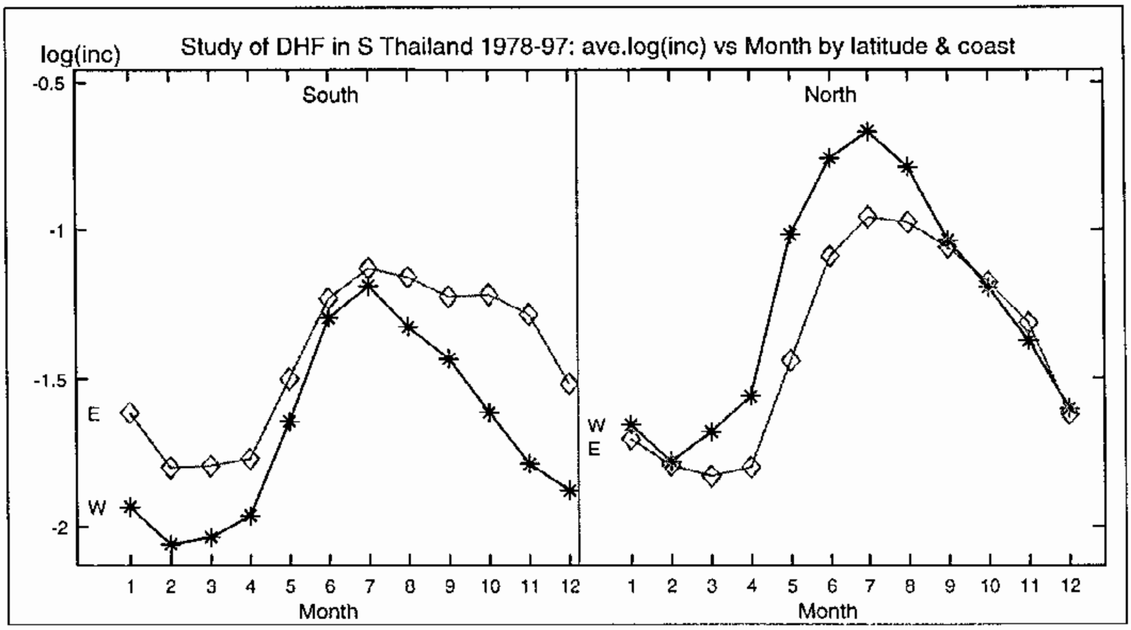
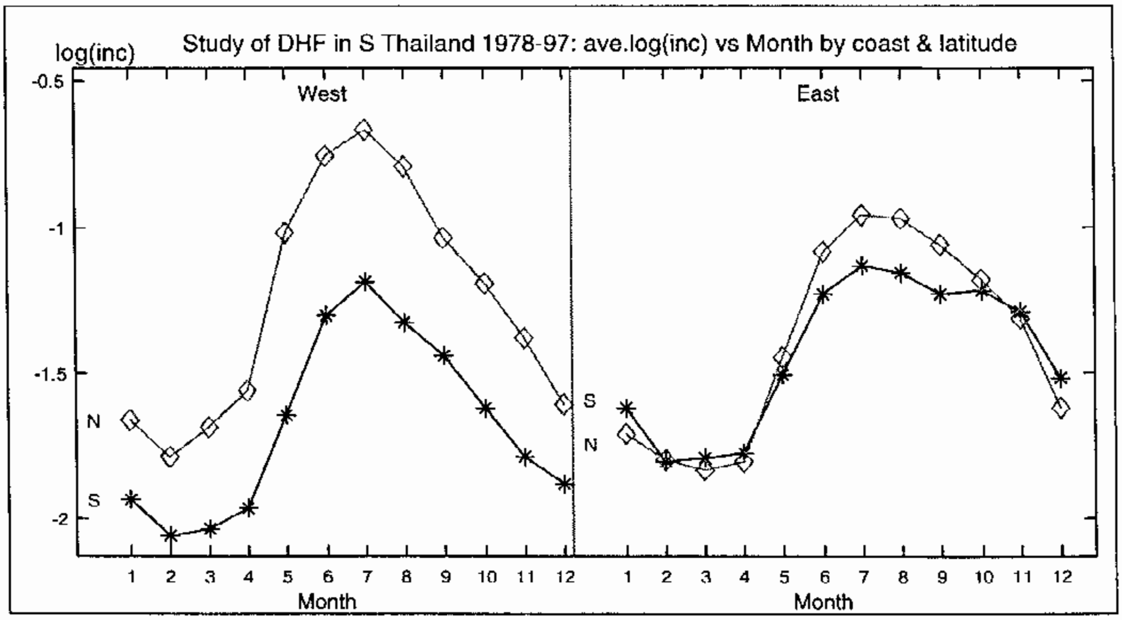


Figure 3.6: The incidence of DHF by month during 1978 – 1997 for the northern and southern locations



These figures clearly show the seasonal pattern of DHF incidence, with the lowest incidences occurring between December and April, and the highest incidences occurring in June, July and August. It is apparent from Figure 3.5 that the DHF incidence in the southern provinces is lower than in the northern provinces. Looking at Figure 3.6, it can be seen that the main difference between the west and east coast provinces is that there is more variation on the west coast, with Krabi having a higher incidence than Trang. There is virtually no difference in the incidence rates for Songkhla and Nakhon Si Thammarat.

The patterns of rainfall over the 20-year period in the four provinces are shown in Figures 3.7 and 3.8. These are computed by averaging the monthly recordings over each year.

From these graphs, it is clear that the data have no trend. In the southern location the pattern of rainfall is not apparently different between the west and east coasts. But in the year 1987 the rainfall between the west and east coasts was abnormal, with Songkhla province having lower rainfall. In the northern location there is more variation between the west and east coasts, Krabi was lower than Nakhon Si Thammarat

Figure 3.7: The rainfall in S Thailand by year (east and west) during 1978 – 1997

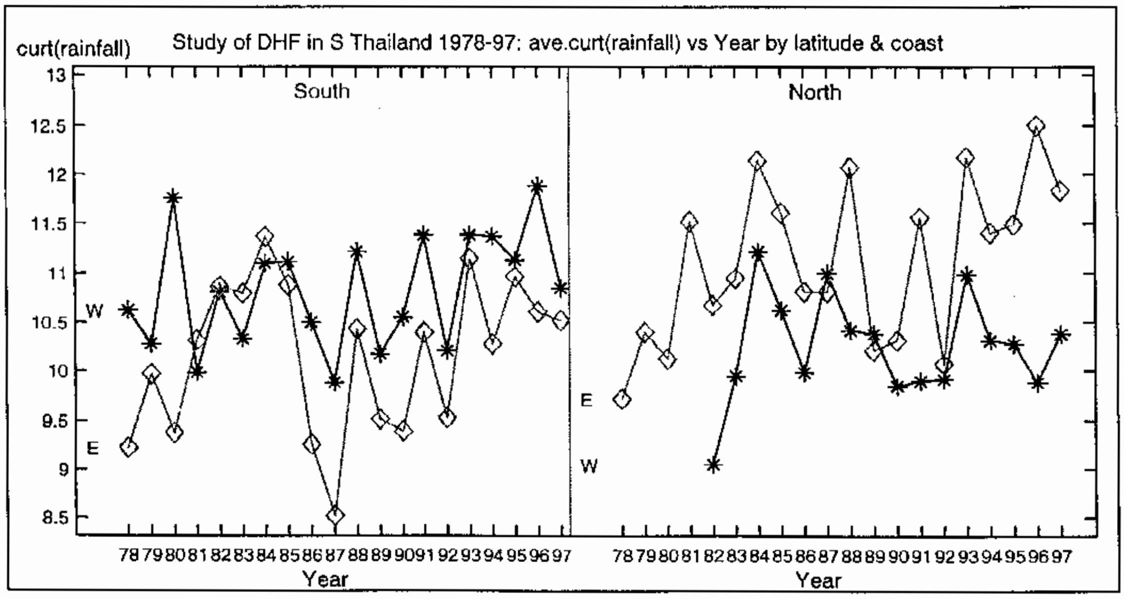
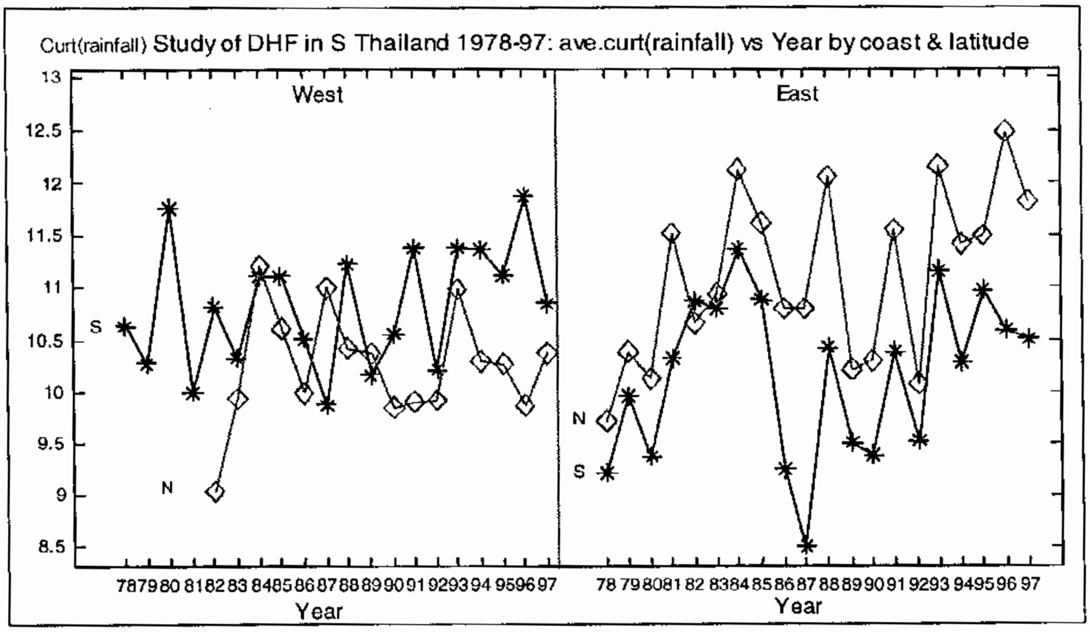


Figure 3.8: The rainfall in S Thailand by year (north and south) during 1978 – 1997



Figures 3.9 and 3.10 show the seasonal patterns of rainfall in the four provinces. The seasonal patterns of rainfall in each month between the west and east coasts, for each pair of provinces are depicted in Figure 3.9. The average rainfall between the west and east coasts is similar. Rainfall is found to be higher in April to November. In addition, the highest peak of rainfall exhibited in the east coast for the southern and northern locations.

The seasonal patterns of rainfall between the southern and northern location for the west and east coasts are similar (Figure 3.10). However, the pattern of rainfall between the west and east coasts is markedly different. The west coast has high rainfall between April to November and low rainfall in December to March, whereas in the east coast highest peak of rainfall is in November.

In the next section the association between DHF incidence and rainfall is investigated, using scatter plots and bivariate time series.

Figure 3.9: The rainfall in S Thailand by month during 1978 – 1997 between the east and west coasts for the southern and northern location

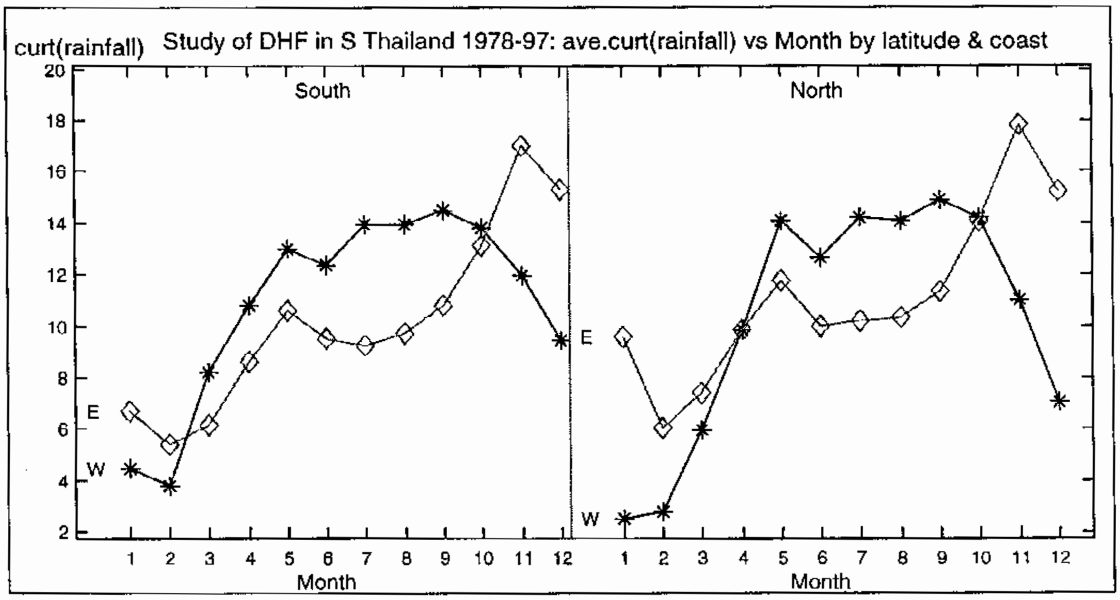
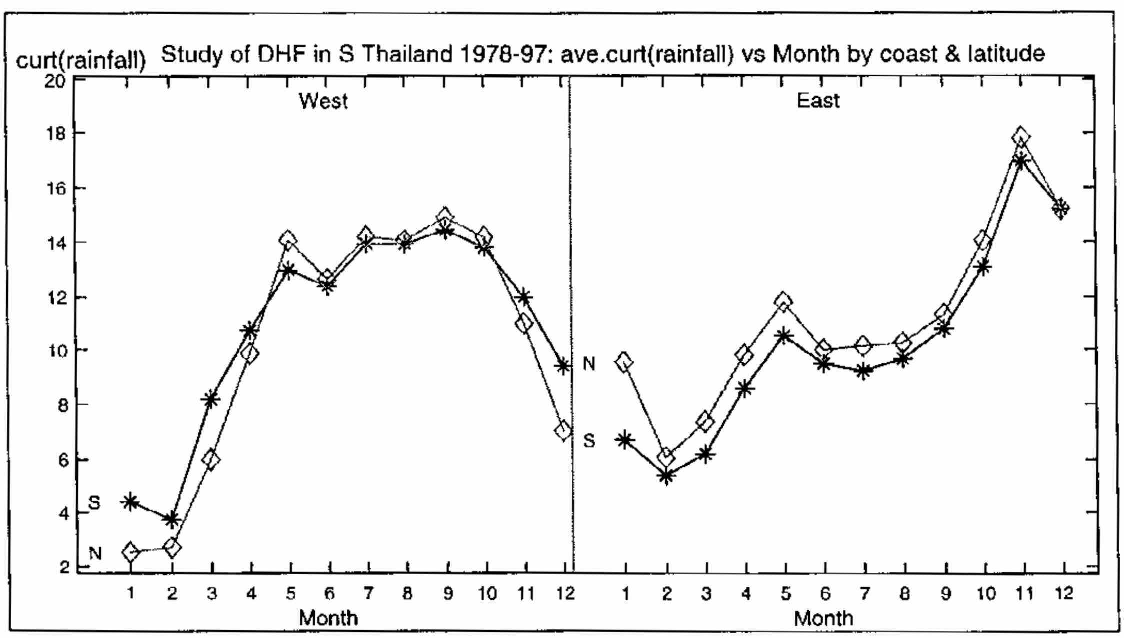


Figure 3.10: The rainfall in S Thailand by month during 1978 – 1997 between the southern and northern location for the west and east coasts



Association between rainfall and DHF incidence

Figure 3.11 shows the association between monthly rainfall and DHF incidence in each of the four provinces over 20-year period. The rainfall strongly correlates with DHF incidence for Krabi (r^2 =0.459) whereas the weakest relationship is found for Nakhon Si Thammarat (r^2 =0.143).

Figure 3.11: The relationship between DHF incidence and rainfall

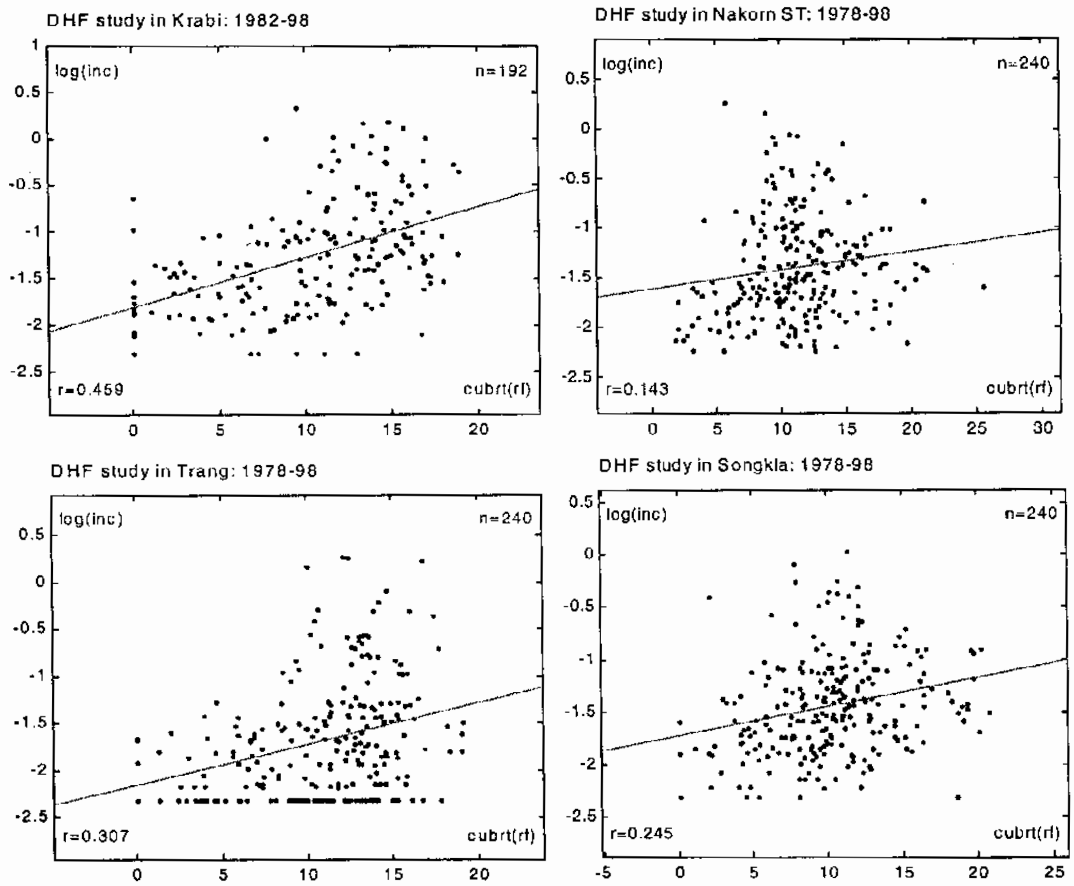


Figure 3.12 shows the DHF incidence and rainfall as time series with aligned horizontal axes, for each of the four provinces.

From this graph, it can be seen that the time series of rainfall in each province is stationary. The stationary patterns of Krabi and Trang are similar, due to their geographic locations on the west coast. Similarly, the patterns of Nakhon Si

Thammarat and Songkhla are similar, due to their geographic locations on the east coast. However, the time series of DHF incidence in each province are not stationary. But DHF incidence in each coast is similar because of the geographical proximity factor.

Figure 3.12: Bivariate time series of DHF and rainfall in S Thailand during 1978-1997