SPATIAL PREDICTION OF MALARIA IN THE RED RIVER BASIN, YUNNAN, CHINA USING GEOGRAPHICAL INFORMATION SYSTEMS AND REMOTE SENSING

Thesis submitted to the University of London for
the degree of Doctor of Philosophy
in the Faculty of Medicine

LUO DA-PENG MB, MSc

Department of Infectious and Tropical Diseases
London School of Hygiene and Tropical Medicine

July 2000

Abstract

This study aims to identify risk factors for malaria related to landscape, environmental, and socio-economic and human behaviour variables in the Red River basin area, Yunnan, China, to develop a predictive model of malaria spatial distribution, and to utilise these to improve malaria surveillance and control in the basin area.

Yunnan is one of the most endemic areas for malaria in China today, particularly in the Red River basin and its border areas. Chloroquine-resistant falciparum malaria is continuing to increase, partly due to immigration and socio-economic development for agriculture in the region. Traditional intensive surveillance systems for malaria are becoming unreliable. The terrain shows considerable variation and some of it is relatively inaccessible. The environment, particularly its land use pattern keeps changing. The present study has used geographical information systems (GIS) and satellite imagery data to identify malaria risk factors related to landscape and environmental data and to develop a predictive model in hope of predicting the risk of malaria transmission and outbreaks and guiding malaria control in the Red River basin area.

The work is in two phases:

Phase I was a retrospective ecological study. It used basic GIS techniques to analyse routine malaria data, and existing environmental and ecological data in the Red River basin area primarily to identify major determinants of malaria spatial distribution related to landscape and environmental variables. The malaria data were those were reported from villages through health care systems. In view of their limitations of accuracy and coverage, phase II was undertaken. However, The work of the phase I study helped to formulate the specific hypotheses for phase II study.

Phase II study was a prospective study. Malaria incidence data were collected in a field survey of the whole study population by our own research team. Malaria incidence data were integrated with altitude data derived from terrain maps and a land use map derived from SPOT 4 imagery data into the GIS. Multilevel Poisson

regression modelling were used to model landscape and land use determinants of malaria.

The phase I study was carried out in one county of the Red River basin area, Yuanyang County, Yunnan, China. Malaria and population data at the level of administrative village were collected in 131 administrative villages of the county. Terrain maps, the land use map, soil map and administrative boundaries of Yuanyang as well as malaria risk maps were integrated within GIS. The data were analysed by modelling the risk of malaria in the administrative villages and their landscape and environmental variables generated from GIS. The results of the analysis revealed that spatial distribution of malaria was determined by the landscape and land use patterns in the administrative villages. Malaria was negatively correlated with mean altitudes of administrative villages, but more paddy and forest would increase the risk of malaria in the villages.

Phase II study was carried out in the Feng Chun Ling Township of Yuanyang County from May to December 1998. The entire population of 24,280 in 5,007 households was included in the study. Around 14% of the study cohort, mostly from the mountains, however, had a history of temporary migration to the lowlands where malaria is highly endemic during the study period. A total of 649 malaria cases (including 3 mixed infections) were identified in the study cohort during the 7-month period. Of the 649 malaria cases, 400 cases were from the population with a history of temporary migration during study period. The overall risk of people with and without temporary migration history were 118.6 and 12.1 per 1000 persons, respectively, during the 7-month study period. The relative risk of the migrated population against non-migrated population were 9.8, suggesting the migrated population had around 10-fold higher risk of malaria than those of non-migrated population. Only 334 indigenous malaria cases out of all malaria cases (649) in the study were used for further analyses and model building. The risk for *P.vivax* indigenous malaria was 17.8 per 1,000 person years at risk and for *P. falciparum* was 6.9 per 1,000 person years at risk. Malaria data were integrated with household locations identified by Global Position Systems (GPS), a land use map derived from a SPOT 4 image and terrain maps into GIS.

The results of multilevel Poisson regression modelling in phase II study revealed that indigenous malaria is negatively correlated with altitude for *P. vivax* and *P. falciparum*. Paddy and forest would increase the risk of malaria, but the effect of forest on malariareached a plateau at certain level of the forest coverage.The mosquito net is protective for indigenous malaria in the analysis. The protective efficacy for *P. vivax* is 40% and *P. falciparum* 29%, respectively.

In conclusion, malaria transmission in the study area is primarily determined by the environmental variables particularly altitude, paddy and forest in the Red River basin area. But human behaviour such as temporary migration and the use of mosquito nets play a very important role in determining the malaria spatial pattern in the study area. Malaria control and surveillance should focus on the lower altitude areas and the mobile population in the Red River basin area. The overall temporarily migrated population plus population living below 1,200 metres accounted for 44.2% (10734/24280) of total population, but accounted for 85.7% (559/652) of total malaria cases during the study period, while migrant plus those living under 800 metres accounted for 73.6% (480/652) of cases in only 21.2% (5155/24280) of the population Future development of land should be aware of the potential for malaria and other vector borne disease risk arising from expansion of paddy field and deforestation, which will provide breeding sites for mosquitoes, particularly in the middle and lower altitude areas. Subsequently, they might result in malaria and other vector borne disease outbreaks or epidemics. The first priority of malaria control strategy in the immediate future is to encourage local residents and 'downhill' migrants to use mosquito nets and to ensure these are regularly treated with insecticides. Chemoprophylaxis and other control measures need to be explored for the temporarily migrating population in the Red River basin area.

Acknowledgement

This study was made possible because of a research training grant offered to the author by the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, who also offered substantial financial support to the cost of fieldwork.

I wish to thank my friend, Dr Zhang Zaixing, the Director of Yunnan Institute of Malaria Control (YIMC), and Dr Yang Huang, the former Director of YIMC for their kind support during my field work in Yunnan, China.

I want to thank my research team in Yunnan. This thesis would not have been possible without their hard work. I especially thank my field coordinator, field supervisors and data manager Drs Guo Hong-Ping, Gu Yun-An, He Yong-Ching, Li Wen-Xue, Bai Zhi-Min, Bai Zhi-Rong, Zhu Jia-Min in this team.

I also wish sincerely to thank the Bureau of Health in Yuanyang County, and Feng Chun Township Government for supporting the fieldwork. The Feng Chun Ling Township Hospital kindly provided the operational base for the study.

I am grateful to Drs Byron Wood, Louisa Beck, Matt Bobo and Brad Lobitz at the Centre for Health Application of Aerospace Related Technologies (CHARRT), National Aeronautics and Space Agency (NASA), USA. They gave me a critical advice and help on the classification of the satellite images covering my study area. They also kindly supported my training on remote sensing in CHARRT, NASA.

I also thank Mr Chris Grundy for his long term advice and help with my GIS work during this training programme, and Drs Paul Wilkinson and Ben Armstrong gave me critical advice on using random effect and multilevel models to analyses my data.

I am very thankful to Dr Ilona Carneiro for her advice on my statistical work and critically going through the draft of Chapters 4 and 5 of this thesis, and for her very helpful comments and suggestions.

I am very grateful to Dr Jonathan Cox for his critical advice and very helpful discussion on GIS and remote sensing throughout my research training, and for his visiting my field work. I also wish to sincerely thank Dr Roger Webber, who supported and assisted both here in school and in the field work.

My sincere gratitude is due to my supervisor Professor David Bradley. He guided me into this fascinating area of using GIS and remote sensing on malaria research. He helped me to obtain extra funding for my field work. He has actively supported and encouraged me throughout my research training programme.

My deep appreciation for my wife, Huijun and my son, Weisang for all their support and encouragement during this thesis work.

Contents of thesis

List of tables

List of figures

Glossary of abbreviations

AVHRR	Advanced very high resolution radiometer
CCD	Cold cloud day
CEPS	County Epidemic Prevention Station
CI	Confidence interval
CIR	Colour-infrared
DBMS	Database Management System
DEM	Digital elevation model
DGPS	Differential global positioning system
ENSO	El Nino Southern Oscillation
GCP	Ground control points
GIS	Geographical information systems
GPS	Global positioning systems
HPEPS	Prefecture Epidemic Prevention Station
HRR	High resolution radiometer
LST	Land surface temperature
MARA	Mapping malaria risk in Africa initiative
MoH	Ministry of Health
MSS	Multi-spectral scanner
NASA	National Aeronautics and Space Administration
NDVI	Normalised difference vegetation index
NOAA	National Oceanographic and Atmospheric Administration
OR	Odds ratio
PI	Principal investigator
PYAR	Person year at risk
RMS	Root mean square errors
RR	Rate ratio
RS	Remote sensing
SAR	Synthetic aperture radar
SSH	Sea surface height
SST	Sea surface temperature
SPOT	*Systeme Pour l' Observation de la Terre*
TM	Thematic mapper

TMS	Thematic mapper simulator
UTM	Universal Transverse Mercator
WGS 84	World Geodetic System 1984
WHO	World Health Organization
YIMC	Yunnan Institute of Malaria Control

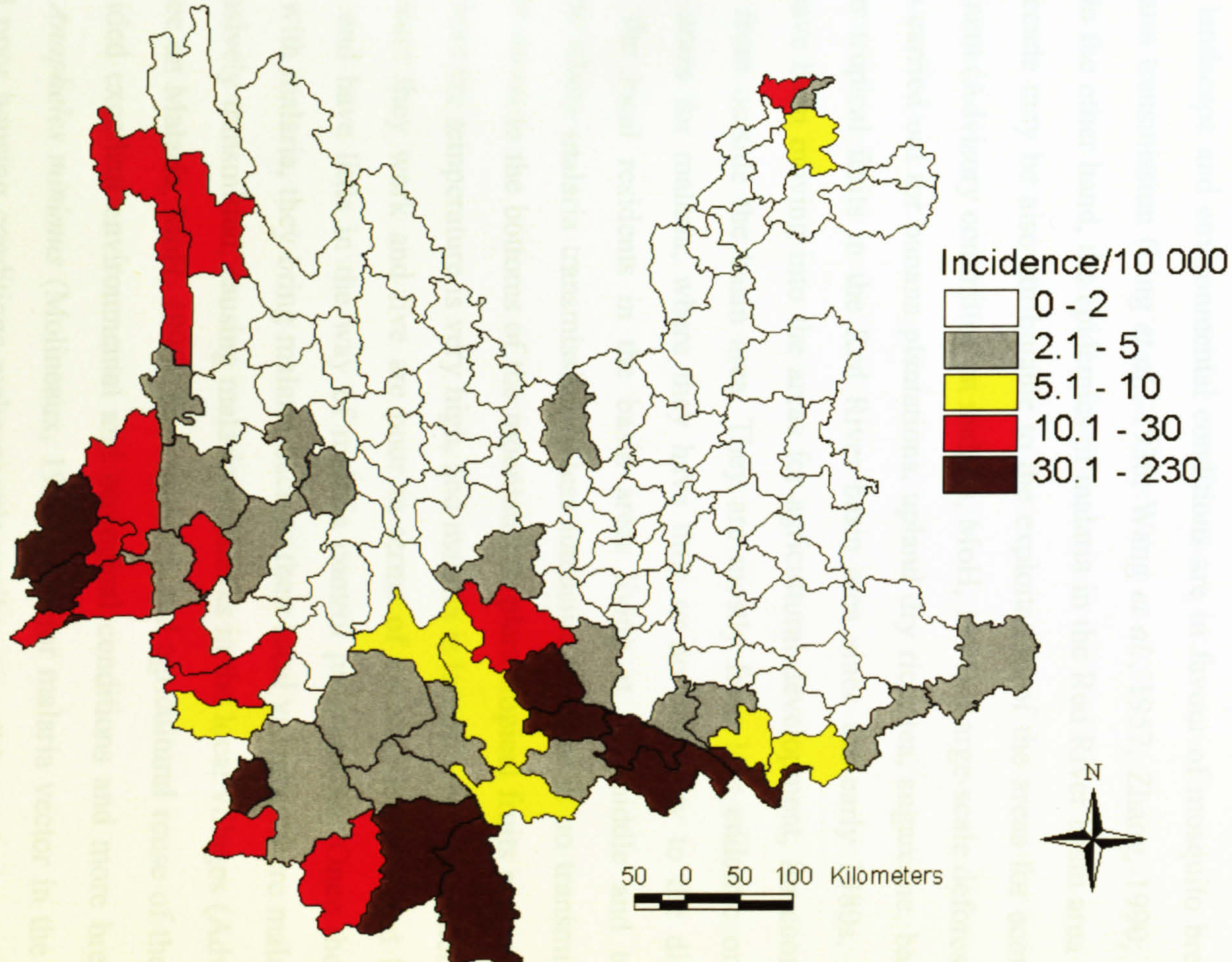
Chapter 1
Introduction

1.1 Rationale and background

Malaria continues to rank as one of the top threats to human beings in the world. Over 2,200 million people are still exposed to malaria risk today. 300 –500 million clinical cases and 1.5-2.7 million malaria deaths occur worldwide each year (WHO, 1993). In some areas, the malaria situation is deteriorating as a result of environmental changes, deforestation, reused land, possible climate change, increasing travelling, increasing drug resistance, and military conflicts and civil unrest. Recent re-assessment of its global malaria control strategy by the World Health Organisation (WHO) pointed to the need for development of more flexible and more reliable malaria control approaches, which considered the requirements of local conditions (WHO, 1993). An essential element in this development is recognising the variability of environmental and epidemiological parameters, which influence malaria transmission. One of the four key technical elements of the strategy is to detect epidemics early so as to contain or prevent them. Geographical information systems (GIS) and remote sensing (RS) might help health planners and policy makers to make more effective control strategies. This thesis describes a study using GIS and remote sensing as the tools to study the malaria situation in the Red River basin area, Yunnan, China, in the hope of finding useful environmental indicators to predict in which areas resources should be concentrated so as to have maximum benefit in the area.

Malaria is still one of most important public health problems in Yunnan Province, China today. Although considerable control of the disease has been achieved in the past three decades, a recent resurgence in malaria has been noted and more fatal cases have been reported in some parts of the province since the middle of the 1980s, particularly for *Plasmodium falciparum* infection (Zhu, 1994). According to the statistical data from the Yunnan Institute of Malaria Control (YIMC) from 1991 to 1995, around 20,000 to 30,000 malaria cases were reported in the province annually. Ten counties with a combined population of 2.5 million in the Red River basin area (the total population of Yunnan was 38 million in 1995) accounted for around 40% of the total malaria cases (35.5%, 37.8%, 44.3%, 45.4% and 38.5% from 1991 to 1995,

Figure 1.1. Malaria distribution in Yunnan, China (1991-1995)



respectively) in Yunnan. Twenty-six counties in the border areas (two counties overlapped with the Red River basin area) with a combined population of about 5 million accounted for about 50% of malaria cases in the province. The spatial distribution of malaria at county level from 1991-1995 is shown in Figure 1.1. Malaria remains highly endemic in the Red River basin area possibly because its natural climate, landscape and environmental conditions are in favour of mosquito breeding and malaria transmission (Ling *et al.*, 1936; Wang *et al.*, 1957; Zhang, 1990; Zhao, 1993). On the other hand, the epidemics of malaria in the Red River basin area in the recent decade may be also attributable to the exploitation of the areas for economic development (Advisory committee on malaria, MoH, 1994). Large-scale deforestation has been carried out for banana plantations, upland dry rice, tea, sugarcane, bananas and other tropical fruits in the Red River basin area since the early 1980s. Many people have been moving into the areas for agricultural development, for economic reasons, from outside the basin area. They are mostly from less endemic or non-endemic areas for malaria, where they have little or no immunity to the disease. Besides, the local residents in the basin areas living at the middle and top of mountains where malaria transmission is less intensive or there is no transmission, also come down to the bottoms of the mountains to plant tropical fruits and crops in places where the temperature is very high, and malaria is intensively transmitted. The places where they work and live are poor in terms of communications and health services, and have little in the way of malaria control programmes. Once they are infected with malaria, they bring malaria back to their local villages where malaria is less intensively transmitted, causing malaria outbreaks in the local villages (Advisory Committee on Malaria, MoH, 1995). Deforestation and agricultural reuse of the land has provided excellent environmental and ecological conditions and more breeding sites for *Anopheles minimus* (Molineaux, 1988), a major malaria vector in the basin area, and poor housing conditions make people easily accessible to the mosquitoes. Nevertheless, the commercial exploitation of the zone in the lower elevation areas remains a top priority policy now for the local government because of over-population at the middle and top of the mountains in the areas. More land will be reused to plant crops and tropical fruits in the near future, which will increase pressure on present malaria control programmes in the Red River basin area.

Chloroquine continues to be used as the first line antimalarial drug in the Red River basin area although it ceased to be used in the border areas in the early 1980s because of the increasing resistance of *P. falciparum* to it (Che *et al.*, 1986). Today, there is increasing concern regarding the effectiveness of chloroquine in the Red River basin area, with some practitioners reporting increasing treatment failures, while others, for reasons unrelated to patient welfare, no longer prescribe chloroquine. A study concerning sensitivity of *P. falciparum in vitro* to available antimalarial drugs was carried out in the basin area in 1993 (Yang *et al.*, 1994), and indicated that the proportion of *P. falciparum* resistant to chloroquine and piperaquine was 78.9% and 72.9% respectively. From 1995 to 1996, a further *in vivo* study revealed that the proportion of *P. falciparum* resistant to chloroquine was around 90.5% in the Manhao Town of the basin area (Guo *et al.,* 1998). *P. falciparum* resistant to chloroquine may contribute to the increase in fatal cases of malaria among local residents and immigrants in the basin area in recent years (Guo Hong-Ping, personal communication).

The distribution of malaria is not homogeneous in the Red River basin area. There are considerable variations of malaria incidence at different altitudes (Ling *et al.*, 1936; Williams, 1941; Wang *et al.*, 1957; Zhang, 1990; Zhao, 1993). Malaria is mainly distributed in the lower altitude areas, and the endemicity of malaria is less intensive in the areas with altitudes over 1,200 metres in the area (Zhu, 1994). Agriculture in the lower altitude areas tends to increase malaria transmission and now leads to epidemics (Zhu *et al.*, 1994; Xu & Liu 1997). Most immigrants clustered in the valley to plant tropical crops and fruits and this resulted in focal malaria epidemics in the valley (Advisory committee on malaria, MoH, 1993).

Funding is one of main problems of malaria control in Yunnan today (Yang *et al.*, 1999). Due to inflation and increasing prices of anti-malarial drugs and insecticides, malaria control programmes have contracted, particularly in remote, newly exploited areas in the Red River basin area. In reality, malaria control has for some time been slack and existing reporting and surveillance systems are poor because of inadequate funding and privatisation of some local health facilities, and local or village doctors who are generally unwilling to report malaria cases to the malaria control authorities. According to the Advisory Committee on Malaria of the Ministry of Health (MoH), the actual number of malaria cases was about 5 - 6 times more than were reported in Yunnan (Advisory Committee on Malaria, MoH, 1993; 1994; 1995). Most unreported cases were falciparum malaria. This situation is apparently even worse in the highly endemic areas, such as the Red River basin and border areas. A recent survey revealed that the actual number of malaria cases was 10 times more than had been reported in the areas (Yang Huang, personal communication). Current malaria stratification systems simply classified all the Red River basin and border areas in Yunnan as "highly endemic" (Zhou, 1981; Zhang, 1991; Liu *et al.*, 1995) although considerable variation of malaria transmission exists in the areas. This might result in malaria control measures being applied indiscriminately or simply no control measures at all because of insufficient funds to cover so large an area. In reality, malaria control measures are not carried out until an outbreak is reported in a place in the Red River basin area . Therefore, an effective predictive system for the Red River basin area is urgently required to enable health authorities to use the limited funds available for malaria control in the areas of relatively highest malaria risk.

Geographical information systems are computer hardware and software packages concerned with the capture, storage, manipulation, query, display and analysis of any and all types of geographical information (Burrough, 1986; Aronoff, 1989; DeMers, 1997). Geographical information plays an important role in the study of disease epidemiology. A study by John Snow (1849) to identify the source of cholera in the nineteenth century is often cited to show the importance of spatial distribution to understanding of diseases. Burkitt (1962) published the geographical distributions of the cases of malignant lymphoma in Africa that led to the hypothesis that an insect vector was important in the aetiology. The principle of using geographical information in disease epidemiological studies, especially for vector-borne diseases is, according to Pavlovsky's classic concept of 'landscape epidemiology' and the 'doctrine of nidality' (Pavlovsky 1945 & 1966), that a disease has its natural habitats in well-defined ecosystems where pathogen, vectors, and natural hosts form associations, within which the pathogen circulates. Therefore, the distribution of a disease is limited by a variety of factors related to its environment, host, socio-economic conditions and vectors, and vectors are also affected by the environment. Both the factors and the sites of infection and disease can be geographically defined. However, quantitative evaluation in epidemiology and ecology had few methods for handing spatial geographical data readily, so that such aspects were either neglected or handled in a crude way prior to adequate computing facilities. Geographical information systems and remote sensing technology should facilitate work in these areas (Bradley, 1994 & 1999).

Remote sensing provides the capability to collect uniform measurements in digital form over a large area at high speed (Aronoff, 1989; Mather, 1999). The data derived from remote sensing such as land use pattern, land cover and meteorological surrogate data, are particularly useful for malaria epidemiology and study of its vectors as has been extensively reviewed (Hugh-Jones, 1989; Washino & Wood, 1994; Hay *et al.*, 1996; Thomson *et al.*, 1996; Hay *et al.*, 1998a). The power of GIS is its ability to effectively integrate various sources and types of geo-referenced data, and to easily update, and efficiently manipulate and display the various spatial data from a large area in a single analysis. Various data sources can be incorporated into a GIS, including existing maps, aerial photographs and satellite images as well as censuses, routine data, epidemiological surveillance data, and data from other specific studies. Remote sensing data can be a theme or a layer in a GIS. Nevertheless, most available geographical and environmental data resources are not generally used in the health sector although they are often used in other fields (e.g. agriculture, water agencies, meteorological organisations, and soil and forest management authorities). In reality within health, GIS and remote sensing have primarily been used to study vector-borne diseases (Hugh-Jones,1989; Washino & Wood, 1994; Mott *et al.*, 1995; Openshaw, 1996; Hay *et al.*, 1996). However, previous studies mostly concentrated on either mapping vector breeding sites based on landscape features, vegetation, soil in a relatively detailed scale, or predicting vector abundance and population dynamics in relatively large areas by using macro-environmental factors such as temperature, rainfall, normalised difference vegetation indexes (NDVI), greenness and wetness, cold cloud duration (Linthicum *et al.*, 1987; Rogers & Randolph, 1991; Wood *et al.*, 1992; Beck *et al.*, 1992; Randolph, 1993; Beck *et al*., 1994 ; Kitron *et al.*, 1994; Roberts *et al.*, 1996; Rejmankova *et al.*, 1996; Thomson *et al.*, 1996; Hay *et al.*, 1996; Cross *et al.*, 1996; Sharma *et al.*, 1996; Beck *et al.*, 1997; Gleiser *et al.*, 1997; Glass *et al.*, 1995; Robinson *et al.*, 1997a &1997b; Dister *et al.*, 1997; Lobitz *et al.*, 2000). Detailed literature reviews concerning GIS and remote sensing and their use in studying malaria and its vectors, will be described in Chapter three. Some studies have used rather simple models and there is a scarcity of work verified by field data sets. There have been few attempts to use GIS to predict the occurrence of vector-borne diseases based on environmental factors and remotely sensed data (Kitron *et al.*, 1994; Malone *et al.*, 1994; Glass *et al.*, 1995; Nicholson & Mather, 1996; Kitron *et al.*, 1997; Malone *et al.*, 1997; Hays *et al.*, 1998b). More recently, people used GIS and remotely sensed data to map malaria and other vector-borne diseases by simple modelling (Craig *et al.*, 1999; Snow *et al.*, 1999a; Snow *et al.*, 1999b; Lindsay & Thomas, 2000). Some studies use GIS to help study malaria epidemiology (Thompson *et al.*, 1997; Binka & Smith, 1998; Hightower *et al.*, 1998; Barrera *et al.*, 1999, Lindsay *et al.*, 2000; Thomas & Lindsay, 2000). The major focus of the present study is to use GIS and remote sensing techniques to directly identify landscape and environmental characteristics and socio-economic and human behavioural factors related to malaria transmission and distribution and to make spatial predictions of the disease.

1.2 General Objectives

The overall aims of the study were to identify risk factors for malaria related to landscape, environmental, ecological and socio-economic and behaviour variables, to develop a predictive model of malaria spatial distribution, and to utilise this to improve malaria surveillance and control in the Red River basin area of Yunnan, China

1.3 General hypothesis

Geographical information systems, integrated with remote sensing and other data sources, can be used to map and identify landscape, environmental features related to malaria distribution and provide reliable spatial predictions of malaria in the Red River basin area of Yunnan, China.

1.4 Strategy of study

In order to test the hypothesis, the following steps have been used in the present study:

Step 1. Phase I study (May - June.1997): to use basic GIS techniques to analyse routine malaria data and existing environmental and ecological data in the Red River basin area, and identify landscape features correlated with the spatial distribution of malaria, and to formulate a preliminary predictive model for the area. The malaria data were those that were reported by villages through health care systems. In view of their limitation of accuracy and coverage, phase II was undertaken. However, the work in phase I study contributed to the formulation of specific hypotheses for the phase II study, which was the main field study.

Step 2. Phase II study (April- Dec.1998): Phase II main field study was a prospective study. Malaria incidence data were integrated with the geographical location of household derived from the Global Positioning System (GPS), altitude data derived from terrain maps and a land use map derived from SPOT 4 imagery into the GIS. Malaria incidence data were collected in a field survey of the whole study population during the main transmission season. The study created a geo-referenced data set of malaria spatial distribution and measured environmental and landscape factors related to the spatial distribution of malaria in the study area, then a predictive model of the study area was developed.

Chapter 2

Malaria and its landscape epidemiology in Yunnan, China

2.1 Geographical and meteorological characteristics in Yunnan

2.1.1 Geography

Yunnan is located in the southwest of China at 97° 31'E - 106 °12' E, 21° 08'N - 29° 15' N with a total area of 394,000 km². It has borders with Myanmar, P.R. Laos and Vietnam in the West and Southwest (Figure 2.1). It shares a border of 4,061 km with the three counties. It is as geographically distinctive as it is diverse. Yunnan is set in the foothills of the Himalayas and directly connects with the plateau of Tibet in the north, faces to India and the Pacific Ocean in the south, and is directly affected by the Southeast and Southwest monsoon.

The terrain of Yunnan consists of mountainous and plateau features. 94% of its land comprises plateau and mountain. The altitude progressively decreases from the Northwest to the Southeast of the province (Figure 2.2). The highest place, Ke Ge Bao Peak in the Mei Li Xie Mountain, is at an altitude of 6,740 metres in the Northwest, and the lowest place is only 76.4 metres above sea level and lies in Haikou County, near the border with Vietnam at the Red River in the Southeast of the province. The terrain can be divided into three strata. Stratum I includes the areas of the Northwest of province, its altitude is around 3,000 to 4,000 metres and most mountains are over 5,000 metres in this stratum. Stratum II includes the middle parts of Yunnan, its altitude is around, 2,300 to 2,600 metres, but the basin areas at the bottom are around 1,700 – 2,000 metres, while the peaks of the mountains usually reach 3,000 – 3,500 metres in this stratum. Stratum III includes the areas of Southern Yunnan and the border areas of the Southeast and Southwest of the province. It consists of hills with an altitude of 1,200 –1,400 metres and of plains and valleys with altitude less than 1,000 metres (Wang, 1990). On average, altitude decreases 6 metres per kilometre from the north to the south. The topographical features of the province are basically plains, valleys, and mountains. People mainly reside in the plains, which only account for 6% of its total land. However, migrant populations and ethnic minority groups tend to live in the valley and mountain areas.

Figure 2.1. The geographical location of Yunnan, China

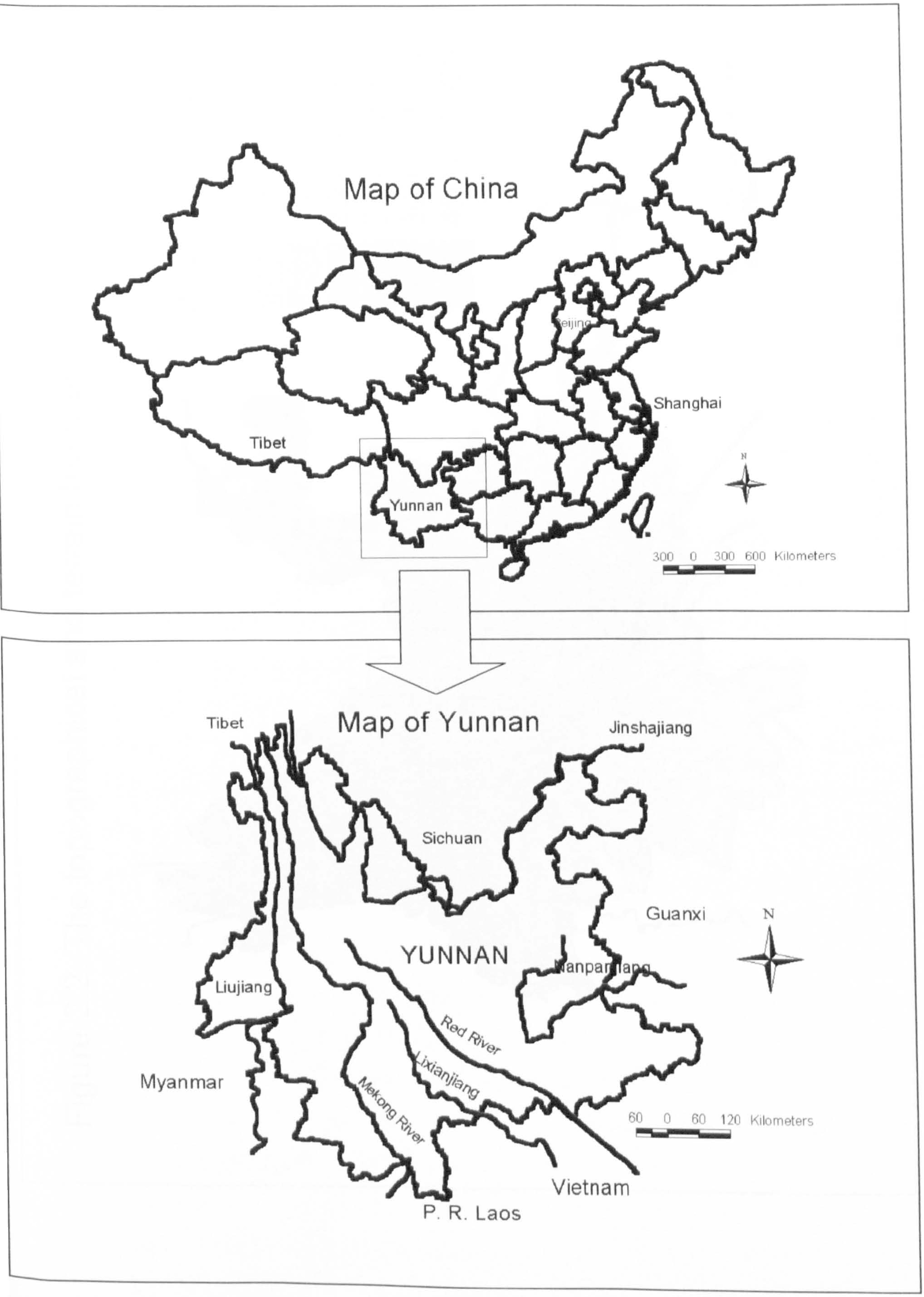
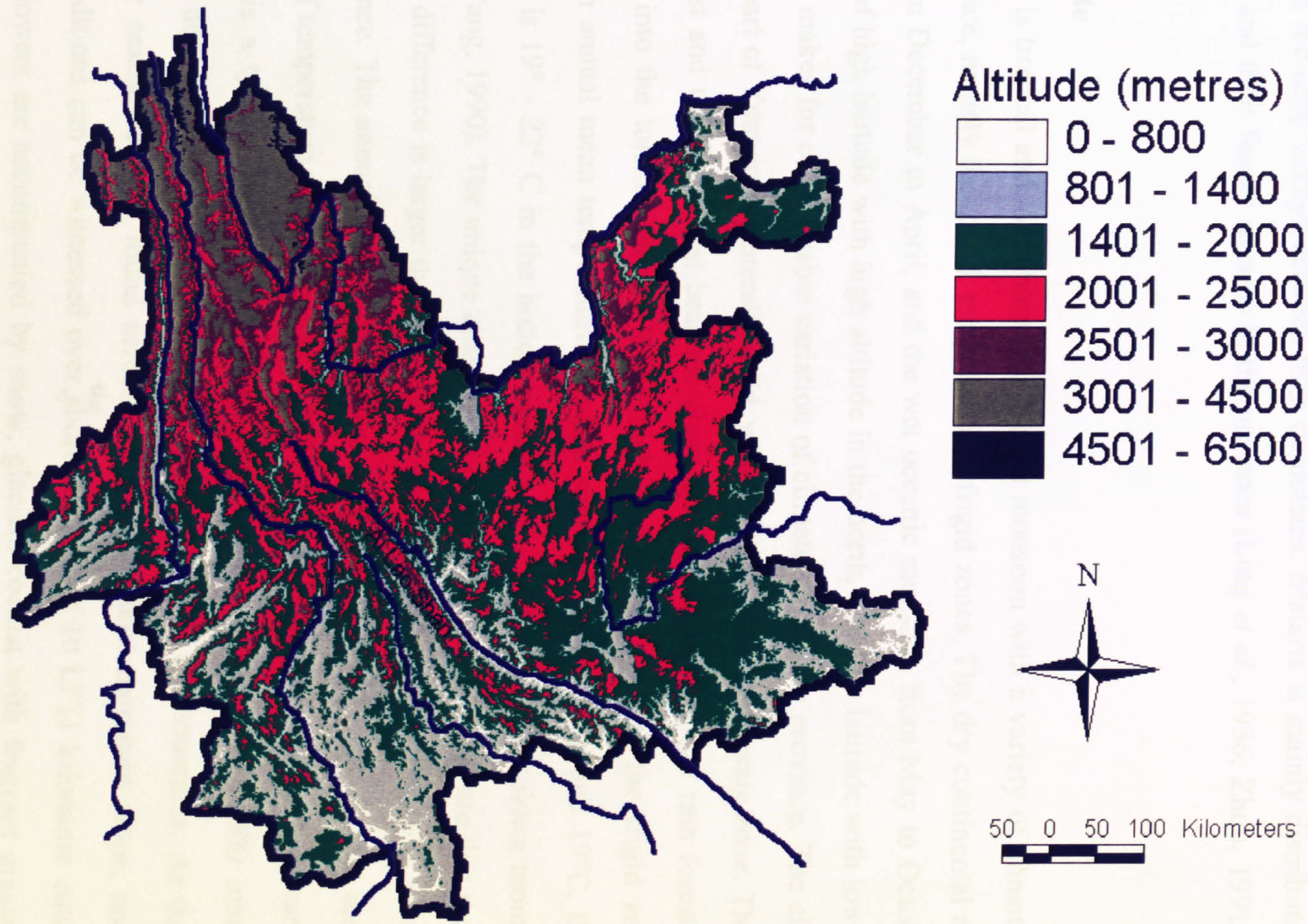


Figure 2.2. The topographical and terrain features in Yunnan, China

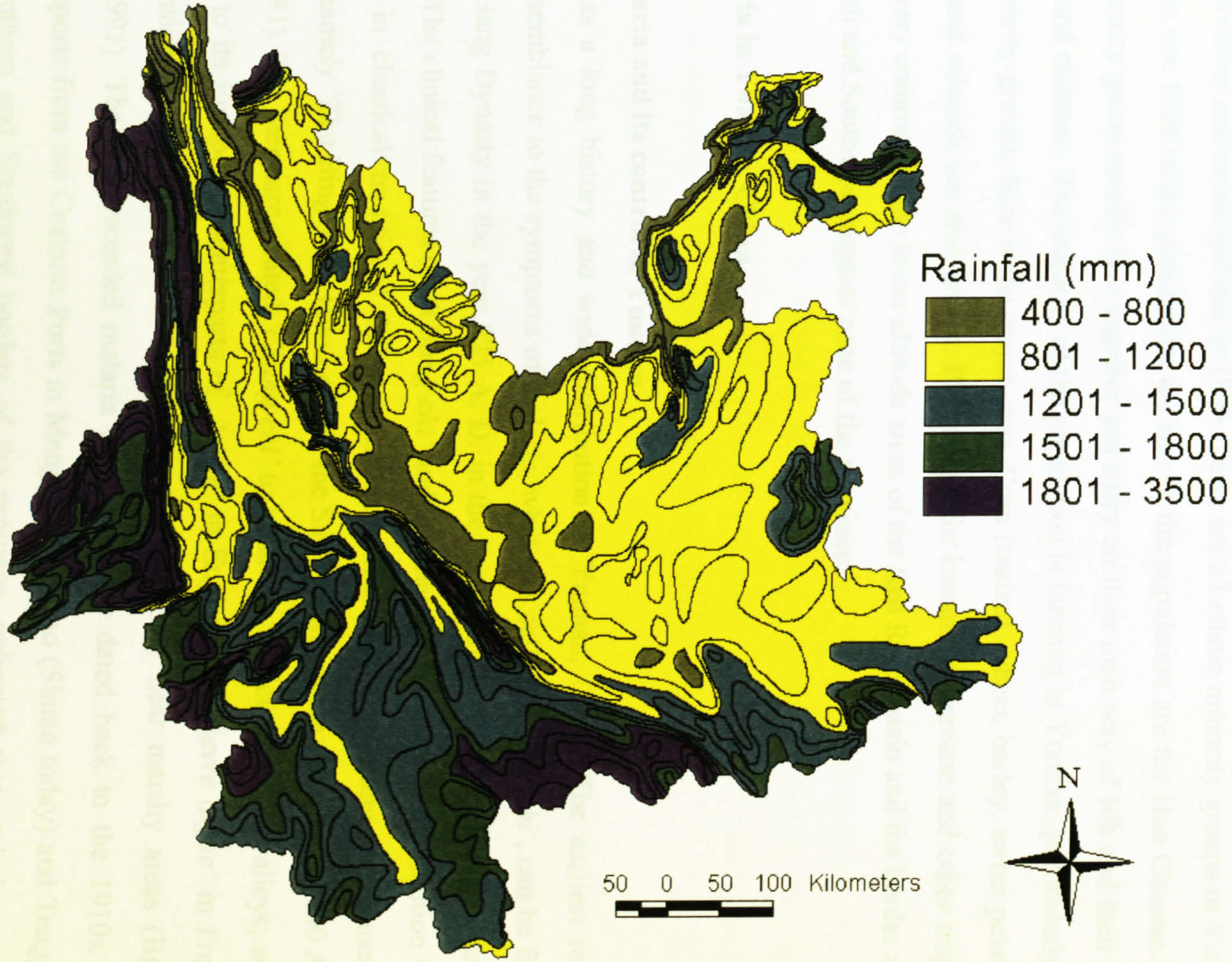


There are plenty of water resources in Yunnan. Six very large river systems (Figures 2.1 and 2.2), namely, the Liujiang, Mekong River, Jinshajiang, Red River, Nanpanjiang, and Lixianjiang rivers, generally run from north to south or to the east. The major rivers usually flow through the deep valleys between high mountains. Malaria is mainly distributed along those rivers and their feeder streams in the province (Ling *et al.*, 1936; Zhou, 1991; Zhu *et al.*, 1994).

2.1.2 Climate

The climate is tropical and /or subtropical plateau monsoon with a variety of climatic zones in the province, namely the torrid, temperate and frigid zones. The dry continental monsoon prevails from December to April, and the wet oceanic monsoon from May to October. The integration of high latitude with high altitude in the north, and low latitude with low altitude in the south makes for considerable variation of climate within the province. The climate in the central part of Yunnan is generally mild, which falls into the temperate zone. The South, the Southeast and the Southwest bear the characteristics of the tropical rain forest climate which falls into the torrid zone while the high Northwest falls into the frigid zone. The difference in annual mean temperature between south and north is around 19°C, the mean temperature is 19° - 22° C in the hottest month and 6° - 8° C in the coldest month in the province (Wang, 1990). The unique characteristics of climate in Yunnan are that the diurnal temperature difference is larger than the annual mean temperature difference in most places in the province. The annual difference of mean temperature is 10° - 12° C, and the diurnal difference of temperature ranges from 12° - 22° C in the same place. The vertical variation of its climate is a typical feature in Yunnan. Most mountains are over 1,000 metres, and temperature decreases 0.5° - 0.7 °C per 100 metres upwards on mountains. As the saying goes " Four seasons can be found across one and the same mountain slope, and varied weather conditions can be witnessed over the short distance of 10 Li"(1 kilometre equals 2 li). Blooming flowers are accompanied by snow, glaciers coexist with fragrant grasses in the same area at different altitudes. The Ai Lao Shan Mountain, as shown in Figure 2.2, defines a line of demarcation of climate in the province. Temperatures to the west of the mountain are 2° C higher than those at the eastside at the same altitude and the same latitude of the mountain in Yunnan (Wang, 1990). The annual mean rainfall is over 1,000 mm in Yunnan, but 85% of the rainfall occurs from May to October. To the South and Southwest there tends to be more rainfall in the province (Figure 2.3).

Figure 2.3. The annual mean rainfall in Yunnan, China



2.1.3 Population and ethnic group

Yunnan is probably the most colourful and most diverse province in China. The particular ethnic mix certainly contributes to this fact. The population of Yunnan was 38 million, spread over 128 counties and municipalities in 1996. There are 25 ethnic minority groups in Yunnan which form one third of the population; the rest of the population are the Han Chinese. The ethnic minority group usually live together and carry on their own way of life and their own traditions and cultures. The majority of the population is farming in Yunnan, particularly the ethnic minority groups. Rice is the major crop. Corn (maize), wheat, barley, sweet potatoes, soybeans and oilseeds are also grown. Plantations for banana, sugarcane and other tropical fruits are very common in the lower altitude areas of the Red River basin and the border areas in the South and Southwest tropical zone of the province.

2.2 Malaria in Yunnan, China

2.2.1 Malaria and its control in Yunnan

Malaria has a long history and wide distribution in Yunnan, China. The earliest record bearing resemblance to the symptoms of malaria, what was called *'Chang chi'*, can be found from the Hang Dynasty in the year 225 A. D. in the book called 'Tai Ping Fan Yiu' in the province. The clinical features of *'Chang chi'* and the characteristics of its distribution were described in classical traditional Chinese medicine literature concerning the causes of diseases, namely 'Zhu Bing Yuan Huo Len', in the Sui Dynasty in the year around 600 A. D. (Zhou, 1991). The literature related *'Chang chi'* to 'mist' or 'malicious air' in valleys, which is similar to its European counterparts of *'ague'* in English and *'fievre aigue'* in French, which considered malaria was caused by foul air commonly near marshy areas (Bruce-Chwatt, 1991). The first recorded malaria cases probably dated back to the 1910s. The medical reports from the Customs Ports in Mengtze, Szchmao (Simao today) and Tengyueh of the Southern and Southwest borders of the province, indicated that malaria cases in Yunnan, especially the 'malignant type' due to *P. falciparum,* were clearly linked to travel for trade to Myanmar (Faust, 1926). Yao and his colleagues carried out their first study on malaria in this province and proved that so called *'Chang chi'* was none other than malaria in Yunnan in 1935 (Ling *et al.*, 1936). Later studies indicated that malaria was widely prevalent in the whole province, especially near the Southwest border areas and the Red River basin area in Yunnan (Ling *et al.*, 1936; Robertson, 1940; Robertson & Chang, 1940; Williams, 1941; Ma, 1942; Yao *et al.*, 1943; Hsu *et al.*, 1949).

There were three well-documented 'malaria' outbreaks in the province in this century. In 1904, about 200,000 to 300,000 workers had been recruited from Yunnan and its neighbour provinces by French colonials to establish a railway from Yunnan to Vietnam. Around 60,000 to 70,000 workers died of 'malaria' during the 7 year period: the so-called 'one sleeper one life' in some places (Zhou, 1991). In 1919, 'malaria' was the cause of an outbreak in Simao, the southern commercial centre in the province. The disease was brought to the city by soldiers, who contracted it from the civilians on the border. The epidemics of the disease had become worse and worse since then. The population of the city had been reduced from 76,800 in 1919 to 24,106 in 1932(Ling *et al.*, 1936) and 944 in 1949 due to 'malaria' (Wang *et al.*, 1957; Zhou, 1991). The third outbreak occurred in Yuanxian in 1933, 30,000 died of malaria in the county within a few years (Zhou, 1991).

In order to determine the distribution, incidence, prevalence and intensity of malaria transmission in Yunnan, province-wide epidemiological surveys were carried out by Chinese scientists in the early 1950's. The results of the surveys revealed that malaria was endemic in 120 out of 128 counties in the province, 65 of which were endemic for *P. falciparum* (Zhu, 1994). Around 140,000 to 420,000 clinical cases per year were reported in Yunnan from 1952 to 1956. However, there were considerable spatial malaria variations: the mean incidence was 249 per10,000 persons per year in the whole province in 1953, but 542 per10,000 persons annually in border areas (Zhou, 1991). Malaria control had started since the end of 1950s, and large-scale DDT residual spraying was carried out in the hyper-endemic areas. Large-scale epidemics of malaria were under control in the whole province soon after the malaria campaign began. Its incidence decreased to 7 per 10,000 persons in the province in 1967. Unfortunately, the malaria control programme was interrupted due to the disbanding of anti-malarial organisations during the upheaval of the 'Cultural Revolution' in the later 1960s. The resurgence of malaria started in 1969. From 1969 to 1973, epidemics occurred in a population of 640,000 in 37 counties and gradually extended to cover 51 counties. 559 people died of malaria during the 4 year period (Zhu, 1994). 154,348 malaria cases were reported in 1973 when the incidence was 57 per 10,000 persons in the province. Consequently, large scale DDT residual spraying, antimalarial chemotherapy in the transmission season and chemoprophylaxis in the peak transmission season were carried out in all main malaria endemic areas in the province. Malaria continued to decline until the early 1980s, and fell to 13 cases per 10,000 in 1979 to 5 cases per 10,000 persons in 1989. However, a resurgence of malaria has been noted especially for *P. falciparum* in the Red River basin and border areas in the recent decade (Zhu, 1994; Zhu *et al.*, 1994; Advisory Committee on Malaria, MoH, 1994). Malaria remains highly endemic in the Red River basin and border areas probably because their climate and environmental conditions are in favour of mosquito breeding and malaria transmission (Ling *et al.*, 1936; Zhang, 1990; Zhao, 1993). On the other hand, the epidemics of malaria in the Red River basin and border areas in recent decades might be also attributable to the exploitation of the areas for economic development and population migration (Advisory Committee on Malaria, MoH, 1994).

2.2.2 Malaria parasites in Yunnan

All four human parasites exist in Yunnan. However, the two epidemiologically important species are *P. vivax* and *P.falciparum*. *P. malariae* and *P. ovale* were detected occasionally.

P. vivax is the most widely distributed malaria parasite in Yunnan. It is occurring throughout the whole province. The local strain of *P.vivax* in the province is similar to the temperate strain with long and short incubation periods, depending on the number of sporozoites inoculated (Yang, 1991). The more sporozoites inoculated, the shorter the incubation period. However, the clinical episodes and relapse patterns of the parasite differ between the southern and northern parts of the province. The febrile episodes have been said by Yang (1991) to occur daily among primary infection cases but to occur every other day among relapse cases in the southern part of Yunnan. The febrile episode occurs every other day for both primary cases and relapse cases in the Northern parts of this province.

P. falciparum is the most harmful parasite and causes the most frequently fatal malaria in Yunnan. The parasite has been diagnosed in 75 counties and was responsible for most epidemics of malaria in the province. It has been mainly distributed to the south of 25° N, particularly in the border areas, the Mekong River basin areas and the middle and lower reaches of the Red River basin area (Zhang, 1990). All endemic areas of *P. falciparum* co-exist with endemic *P. vivax* in Yunnan, but *P. falciparum* is the dominant species in the border areas and lower altitude areas of the Red River basin area. Chloroquine resistant

P.falciparum is increasingly becoming an important medical and public health problem in Yunnan and will be dealt with in section 2.2.4 of this Chapter.

2.2.3 Malaria vectors in Yunnan

Forty-one species of *Anopheles* have been identified in the province since 1935. Six of them, namely *An. minimus, An. sinensis, An. anthropophagus, An. kunmingensis, An. jeyporiensis* and *An. candidiensis*, were proved to be malaria vectors in Yunnan, and the first four anophelines are the most important (Dong, 1993).

An. minimus is a main malaria vector in the province. It was identified in 89 out of 128 counties in Yunnan. The range of its distribution is from 28 °N at Shejian county to 21° N at Mongla county (Dong, 1993). It was first identified by Yao and colleagues (Ling *et al.*, 1936). They dissected 30 *An. minimus* during their study, of which one was found sporozoite-positive in its salivary glands, and they hypothesised that *An. minimus* might be one of main vectors for malaria transmission in the province. This hypothesis was supported by several large-scale studies (Roberston *et al.*, 1940; Yao & Pei, 1943; Chow & Balfour, 1949; Wang *et al.*, 1957). The population density of the mosquito shows considerable variations with different latitude, altitude, vegetation and river systems in the province (Dong, 1993). It is mainly found in areas with altitudes from 600 metres to 1,400 metres, but it can be found in the same area with an altitude up to around 1,600 metres. Based on the statistics of Dong (1993), 80,669 *An. minimus* were dissected from 32 sites in 14 counties in the province from 1935 to 1982. 580 of them were identified as sporozoite positive in their salivary glands from 26 sites in 11 counties. The mean proportion of mosquito sporozoite-positive in their salivary glands was 7.2 per thousand. *An. minimus* is endophilic and endophagic and prefers human blood. It likes to breed in slowly running water such as streams and in clear water bodies at the edge of forest in Yunnan (Dong, 1993). Its breeding places have usually a certain amount of shade. In some places, it might be found in terraced type paddy fields, seepage from springs, and small water reservoirs with clear water in the Southwest part of the province (Dong, 1993).

An. sinensis is the commonest species of *Anopheles* in Yunnan. It is distributed over the whole province, except at an altitude over 4,000 metres in the Northwest part of the province. A total of 49,580 mosquitoes has been dissected so far at 38 sites in 15 counties, 12 of which from 4 counties were identified as sporozoite positive in their salivary glands, giving a proportion positive of 0.24 per thousand mosquitoes in the province (Dong, 1993). Although it is an abundant mosquito, it is not an efficient vector. It prefers animal blood, mainly buffalo . *An. sinensis* mainly breeds in paddy field and ponds. The density of its population is correlated with the areas of paddy fields in the locality.

An. anthropophagus is mainly distributed to the north of 27° N in the Northeast part of this province. Occasionally, it can be found in the western parts of Yunnan. It is an anthrophilic and endophilic mosquito. It mainly breeds in small shaded ditches, but some breed in paddy fields. 1,090 mosquitoes were dissected in Shijian county; 3 of them (2.8 per thousand mosquitoes) were sporozoite positive (Dong, 1993).

An. kunmingensis is mainly distributed to the north of 25° NL, in which area the altitude is over 1,700 metres in the Northwest part of the province. Its breeding places are more or less like those of *An. anthropophagus*. 2 out of 480 mosquitoes were identified sporozoite positive (Dong, 1993)

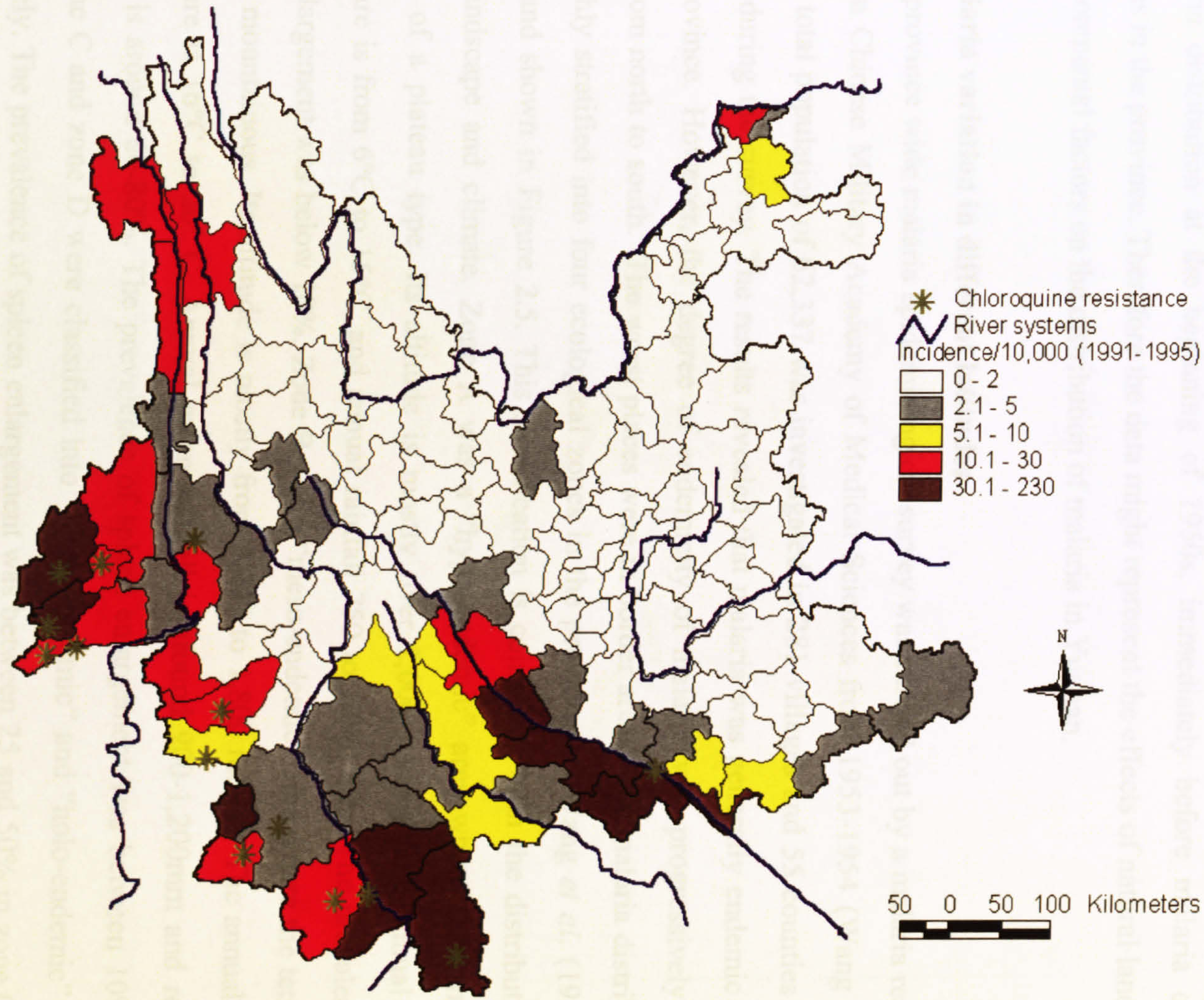
2.2.4 Anti-malarial drug resistance in Yunnan

Resistance of *P. falciparum* to chloroquine was first identified in the Sino-Myanmar border at Genma, Yunnan in 1973 (P.L.A. Kuanmin Medical Research Institute, Malaria Research Group, 1978). 72 cases were studied *in vivo*, 23 of them were classified as S-RI (31.9%), 47 (65.3%) RII, 2 (2.8%) RIII. The strains of *P. falciparum* resistant to chloroquine were identified in all endemic areas in the later province-wide investigation (Che *et al.*, 1986). From 1972 to 1988, 530 falciparum malaria cases were studied *in vivo,* of which 77.6% cases showed resistance to chloroquine; the proportions of RI, RII and RIII were 38.30%, 25.28%, and 13.96%, respectively (Zhu, 1994; Yang *et al.*, 1994). Although most of the *in vivo* studies suffered from multiple problems, such as selection biases, chance of re-infection, loss to follow-up, small sample size, and some information biases , chloroquine resistance had been thought to be well established in most counties in border areas of Yunnan since the early 1980s. The health authority in Yunnan changed from chloroquine to sulfadoxine-pyrimethamine (SDX/PYR) and piperaquine as its first line antimalarial drugs for *P. falciparum* in the border areas in 1983 (Che *et al.*, 1986). The resistance of *P. falciparum* to chloroquine in the border areas continued to be under surveillance, though *in vitro* in 1989 and 1990, the resistance rate was near to 98.7% (75/76) (Yang, 1994).

In recent years, endemic areas of *P. falciparum* have extended from the border areas to the Red River basin area in the province. The sensitivity of *P. falciparum in vitro* to available antimalarial drugs was measured in the Red River basin area in 1993 (Yang, 1994). That study revealed that the strain of *P. falciparum* in the basin area was highly resistant to chloroquine. The proportions of *P. falciparum* resistant to chloroquine and piperaquine were 78.9% and 72.9% respectively. A further study concerning sensitivity of *P. falciparum in vivo* to chloroquine from 1995 and 1996 indicated that the proportion of *P. falciparum* resistant to chloroqine was 78.5% (Guo *et al.*, 1998). The spatial distribution of *P. falciparum* resistant to chloroqine, which has been identified *in vivo*, is shown in Figure 2.4.

The first study on SDX/PYR efficacy on *P. falciparum* in the well-known chloroquine resistant area of Xi Shang Ban Na, Yunnan was by the P. L. A. Kuanmin Medical Research team in 1983 (cited by Zhou, 1991). Sixteen local falciparum malaria cases were tested, 12 (75%) of cases were susceptible to SDX/PYR. In 3 cases, however, falciparum parasites didn't disappear within 7 days (RII); and in one case, the parasite disappeared within 7 days but recrudesced on day 12 (RI). In 1992, the sensitivity of *P.falciparum* to SDX/PYR *in vitro* was determined (Yang *et al.*, 1994) in Mongla near the Sino-Laos border. Results of the study suggested that *P. falciparum* was resistant to SDX/PYR. Therefore, there is multiple drug resistance of falciparum malaria in Yunnan today.

Figure 2.4. Spatial distribution of chloroquine resistance of P. falciparum in Yunnan, China



3 Landscape epidemiology of malaria in Yunnan

The complicated climate and landscape conditions make for considerable variations of the spatial distribution of malaria in the province. The following paragraphs review the studies on malaria distribution at the beginning of 1950s, immediately before malaria control campaigns in the province. Therefore the data might represent the effects of natural landscape and environmental factors on the distribution of malaria in Yunnan.

2.3.1 Malaria variation in different latitudes

The first province wide malaria epidemiological survey was carried out by a malaria research team from Chinese Military Academy of Medical Sciences from 1953-1954 (Wang *et al.*, 1957). A total population of 62,337 was investigated in 281 villages and 55 counties in the province during the survey. The results revealed that malaria was generally endemic in the whole province. However, the degree of endemicity of malaria was progressively more intense from north to south. The worst places were in border areas. The malaria distribution was roughly stratified into four ecological zones in the province by Wang *et al.* (1957) as redrawn and shown in Figure 2.5. This stratification is coincident with the distribution of natural landscape and climate. Zone A was a "hypo-endemic" area where the terrain is generally of a plateau type. Its altitude is mostly over 2,000 metres. The annual mean temperature is from 6°C to 15°C, and annual rainfall 750 to 1,000 mm. The prevalence of spleen enlargement was below 10%. Zone B was a "meso-endemic" area where the terrain is generally mountainous. Its altitude is mostly from 1,200 to 1,800 metres. The annual mean temperature is 16°C to 20° C. Annual mean rainfall is around 900-1,200mm and relative humidity is around 60-80%. The prevalence of spleen enlargement was between 10% and 25%; Zone C and zone D were classified into "hyper-endemic" and "holo-endemic" zones respectively. The prevalence of spleen enlargement was between 25 and 50% in zone C. The prevalence of spleen enlargement was over 50% near border areas (Zone D). The altitudes in Zones C and D, are mostly lower than 1,200 metres, with annual mean temperature over 20°C, annual rainfall over 1,200 mm, and mean relative humidity over 80%. This stratification roughly represents the effect of natural landscape and environmental factors on the risk of malaria in Yunnan. The stratification remains broadly valid today although malaria is becoming recognised as a major problem in the Red River basin area in Yunnan today (Zhu, 1994; Advisory committee on malaria, MoH, 1994). Zhang (1990) summarized all surveys in 49 counties of Yunnan from 1950 to 1957. He also stratified malaria distribution into four zones according to latitudes (Table 2.1). His stratification is roughly similar to early stratification by Wang *et al.* (1957) but emphasises that the Red River should be is an important line of demarcation for malaria stratification in Yunnan.

Figure 2.5. Malaria stratification in Yunnan, China by Wang *et al.*(1957)
(see p. 38 of these for explanation)

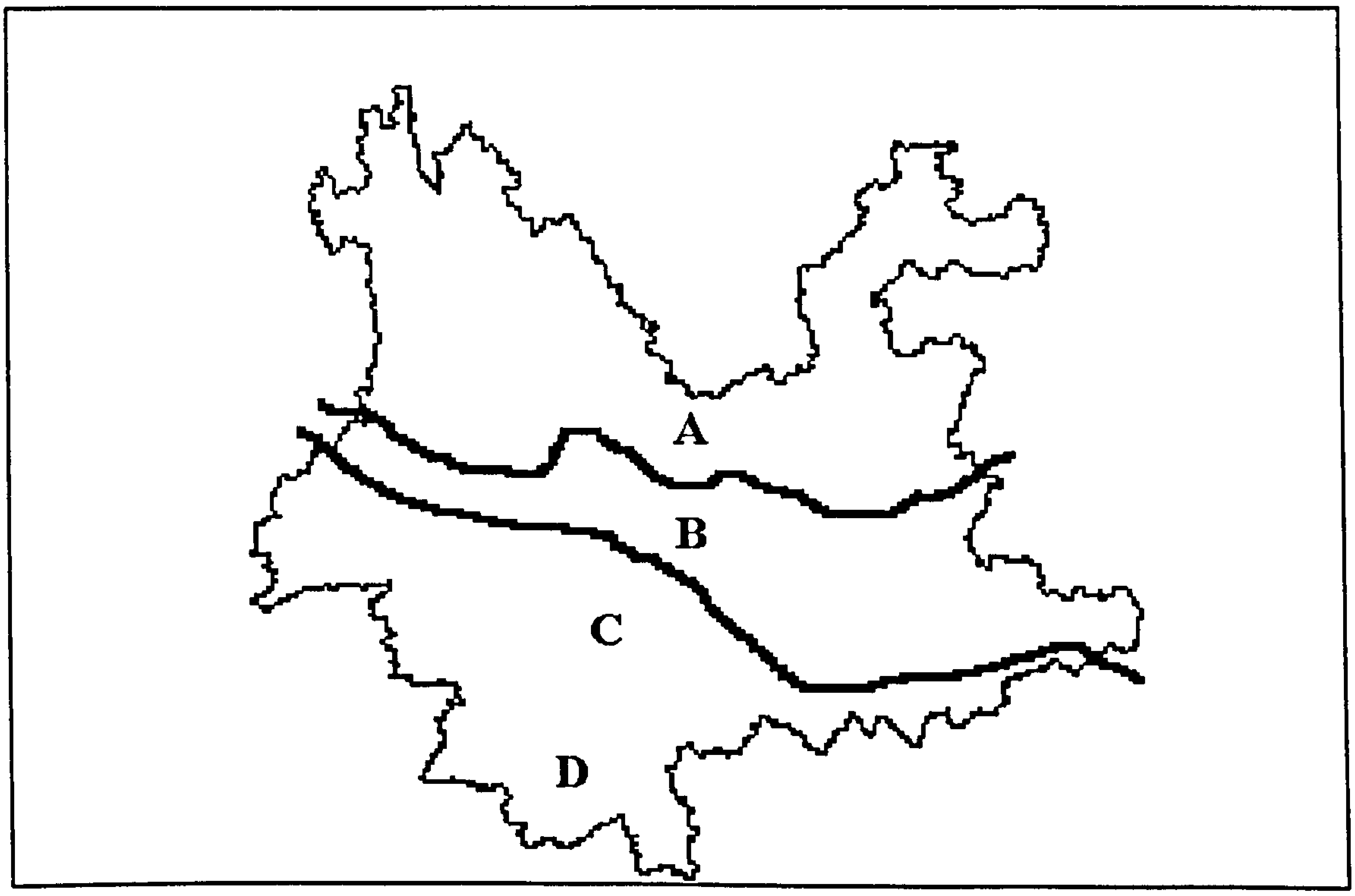


Table 2.1. Malaria variation at different latitudes in Yunnan, China (1950-1957)

	Spleen examination			Malaria parasitaemia		
	No. exam.	No. posi.	%	No. exam.	No. posi.	%
South of 23° 30'NL	120,106	47,910	46.9	79,673	24,450	30.7
23° 30' – 25° NL, West of the Red River	112,958	46,742	41.4	60,624	10,372	17.1
23° 30' – 25° NL, East of the Red River	11,151	2,644	23.7	10,711	722	6.7
North of 25° NL	9,473	2,346	24.8	8,397	716	8.5

Sources of data: Zhang (1990)

2.3.2 Malaria variations at different altitudes

The connection between altitude and intensity of malaria transmission had been made as far back as the first preliminary survey in Yunnan (Ling *et al.*, 1936), and was further quantified by many researchers (Roberston, 1940; Williams, 1941; Wang *et al.*, 1957 & Zhang, 1990). The highest place found with malaria infection was 2,520 metres above sea level in Weixi county in the province (Lian, 1984), slightly lower than 2,800 metres, a world's highest spot where malaria infection was reported (Bruce-Chwatt, 1991). All studies at different altitudes indicated a consistent conclusion that altitude was negatively correlated with prevalence of malaria (Figure 2.6). According to the empirical model by Huo Rong Huo (1984), malaria was stratified into three strata in Yunnan, based on the altitude and malaria incidence data before 1957 at a time when the malaria control programme had not yet begun, namely, "hyper-endemic", "meso-endemic" and "low-endemic" areas. The "hyper-endemic" area had an altitude under 1,200-1,300 metres, the "meso-endemic" area was at an altitude from 1,300 – 1,800 metres and "low-endemic" area with elevation over 1,800 metres. The ranges of altitudes might have some variation between southern and northern part of the province. In the "hyper-endemic" area, the mean temperature is over 18°C and the malaria incidence usually over 20% annually, spleen rate over 25%, *An. minimus* is the dominant vector in this area, more than 20% of total malaria cases were *P. falciparum* infection. In the "meso-endemic" area, annual mean temperature is around 16-18°C, the malaria incidence rate was usually less than 20% annually, spleen rate less than 20%, *An. minimus* and *An. sinensis* are malaria vectors in the area. The proportion of *P. falciparum* infection over total malaria cases in the areas was less than 20%. In "low-endemic" area, the annual mean temperature is less than 16-18°C, malaria incidence were usually less than 5% annually, spleen rate less than 5%, *An. sinensis* is a malaria vector in this area, and no *P. falciparum* infection occurred in the area. This stratification is much more rational and accurate than that of the first stratification by Wang *et al.* (1957) and later re-analysis by Zhang (1990). But it would be very difficult to draw on a map that can be visually appreciated by the malaria control authorities prior to GIS. We use digital elevation model (DEM) to stratify the malaria in Yunnan based on the Huo Rong Huo empirical model, as shown in Figure 2.7.

Figure 2.6.The prevalence of parasitaemia and spleen enlargement at different altitudes in the selected counties, Yunnan (Zhang, 1990)

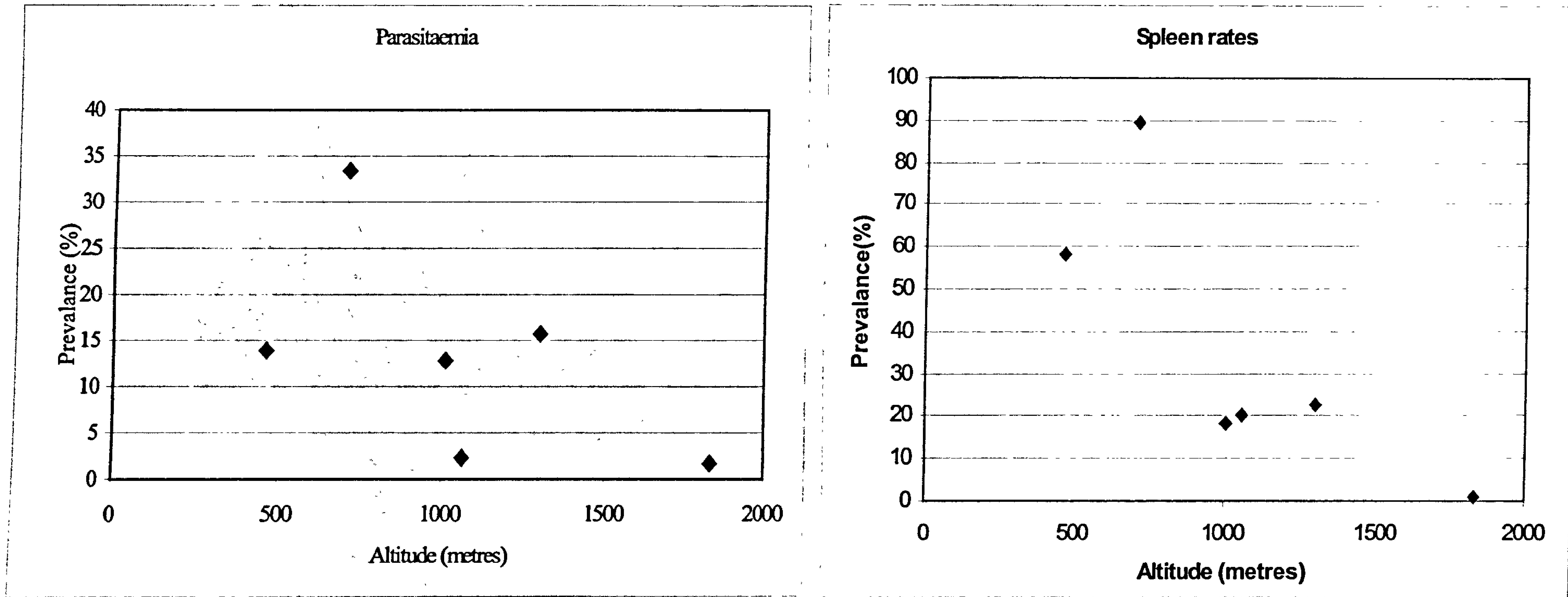
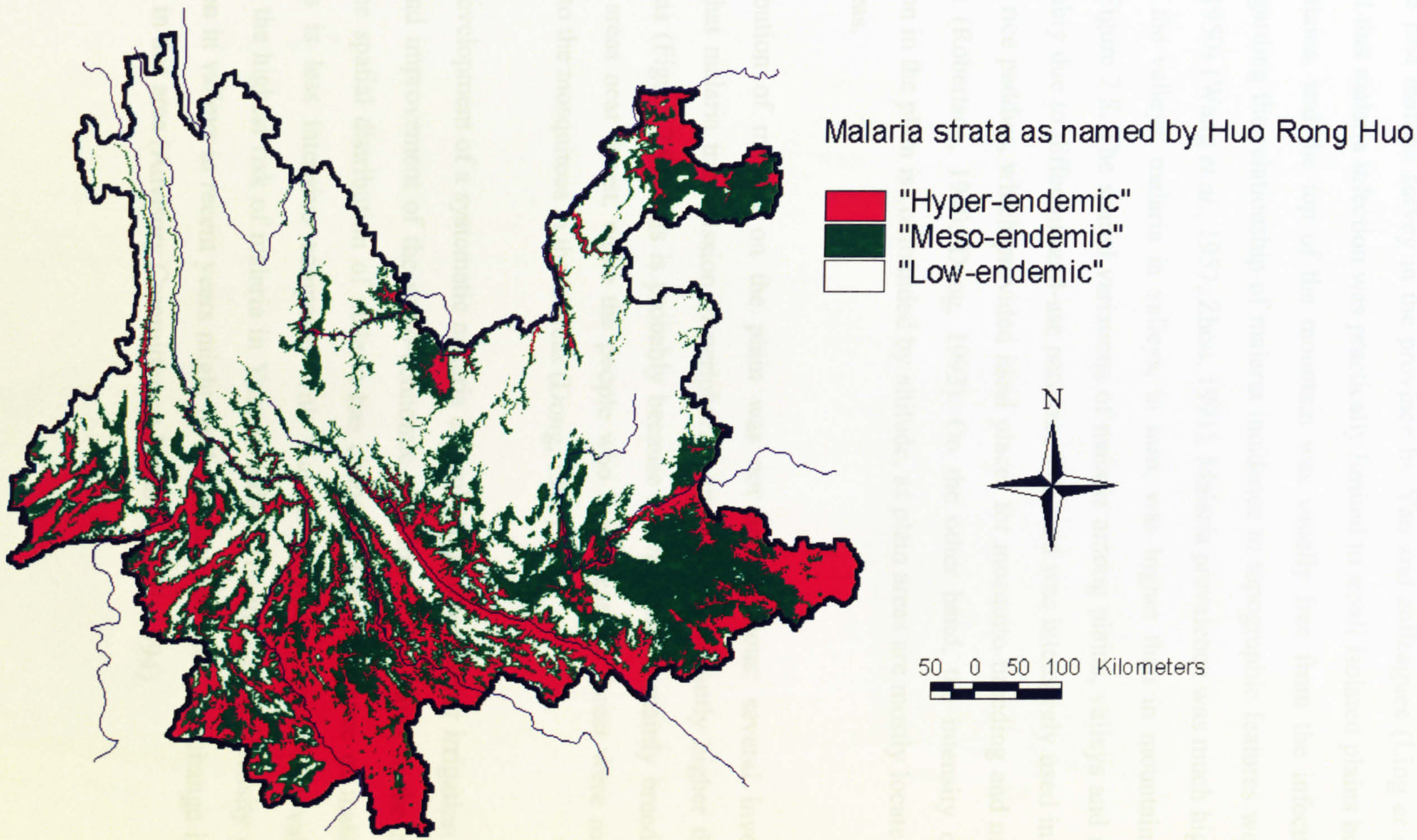


Figure 2.7. Malaria stratification based on the altitude criteria of Huo Rong Huo (1984)



2.3.4 Malaria variation in different topographic features

The first study that noted malaria variation with different topographic features probably dated back to the first malaria survey in the province by Yao and colleagues (Ling *et al.,* 1936). They noted that malaria infection was practically limited to small isolated plains bounded by high mountains, and the top of the mountain was usually free from the infection. More surveys regarding the relationship of malaria incidence to topographic features were carried out in the 1950s (Wang *et al.*, 1957; Zhou, 1991). Malaria prevalence was much higher in the plain than the valleys; malaria in valleys, in turn, was higher than in mountain areas, as shown in Figure 2.8. The spatial variations of malaria among plains, valleys and mountains were probably due to different land-use patterns. The land was intensively used in the plain, mostly for rice paddies, which provided ideal places for mosquito breeding and multiplying in Yunnan (Robertson, 1940; Dong, 1993). On the other hand, the intensity of malaria transmission in the plain is confounded by altitude, as plain areas are mostly located at lower altitude areas.

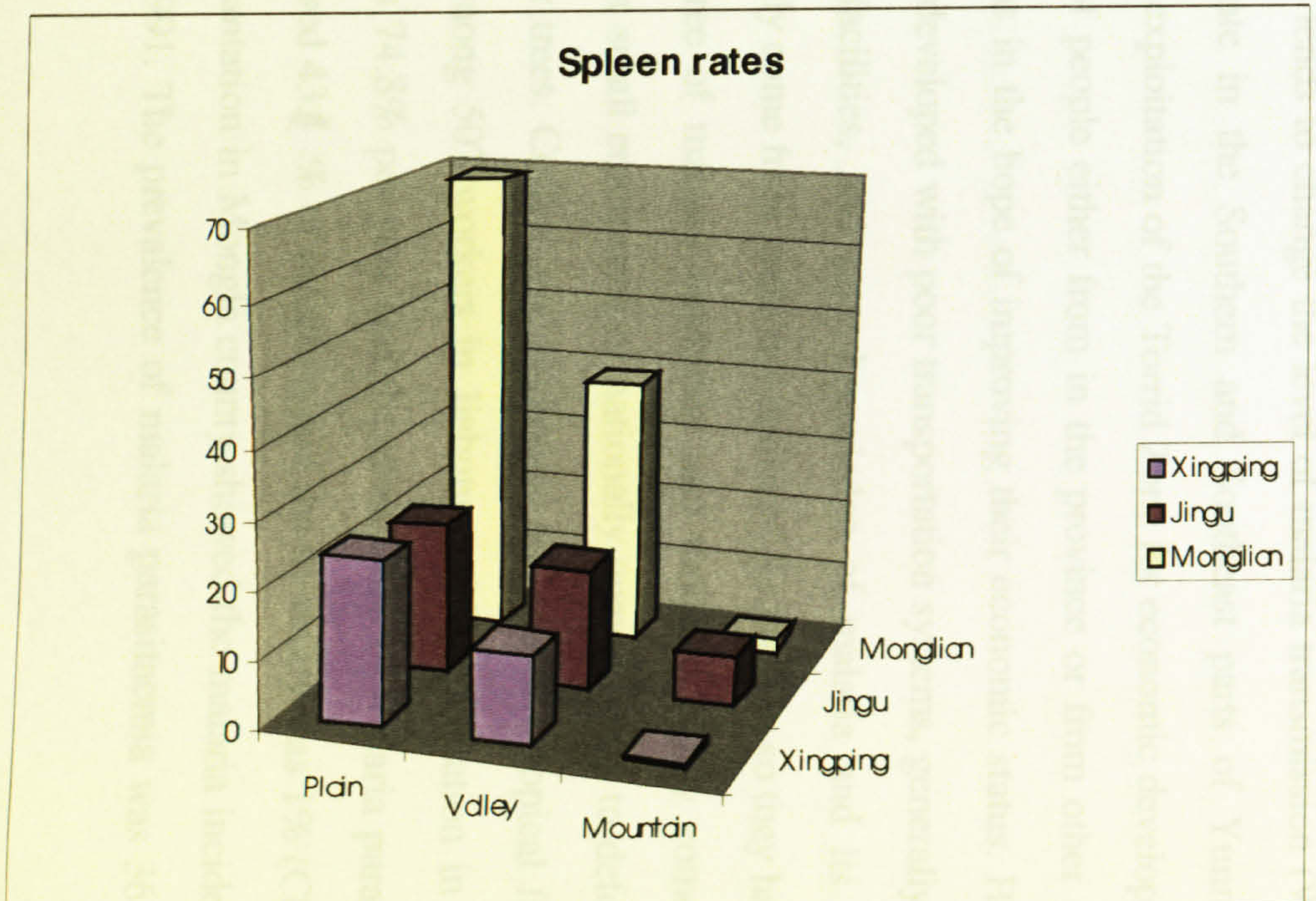
The distribution of malaria on the plain was not homogeneous: several investigations indicated that malaria transmission in peripheral parts was significantly higher than in the central areas (Figure 2.9). This is probably because *An. minimus* is mainly breeding in the peripheral areas near forest, while the people who lived near the forest were more easily accessible to the mosquitoes in those areas (Dong, 1993).

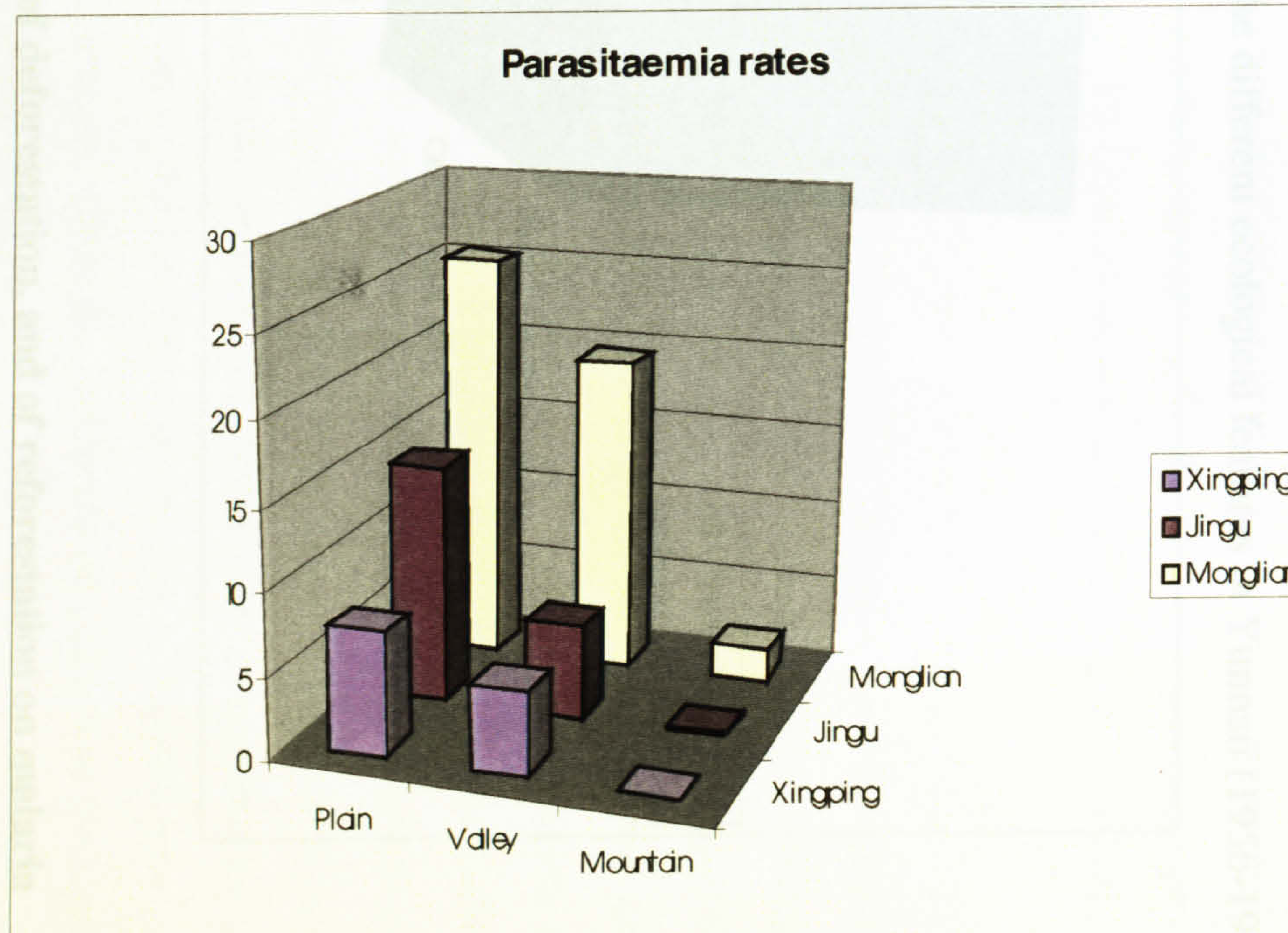
With the development of a systematic malaria control programme, better irrigation and canal systems, and improvement of the socio-economic situation in the plains in the recent two decades, the spatial distribution of malaria has changed. The prevalence of malaria in the plain areas is less intensive compared with that in mountain areas. The valley areas experience the highest risk of malaria in Yunnan today. The increasing intensity of malaria transmission in valleys in recent years might be due to deforestation and change in the land use pattern in this area (Advisory Committee on Malaria, MoH, 1994).

Figure 2.8 Malaria variation in different topographical features in Yunnan, China (1956-1957)

	Spleen rate(%)		
County	*Plain*	*Valley*	*Mountain*
Xingping	24.0	13.1	0.6
Jingu	22.3	17.8	7.6
Monglian	69.5	39.8	2.0

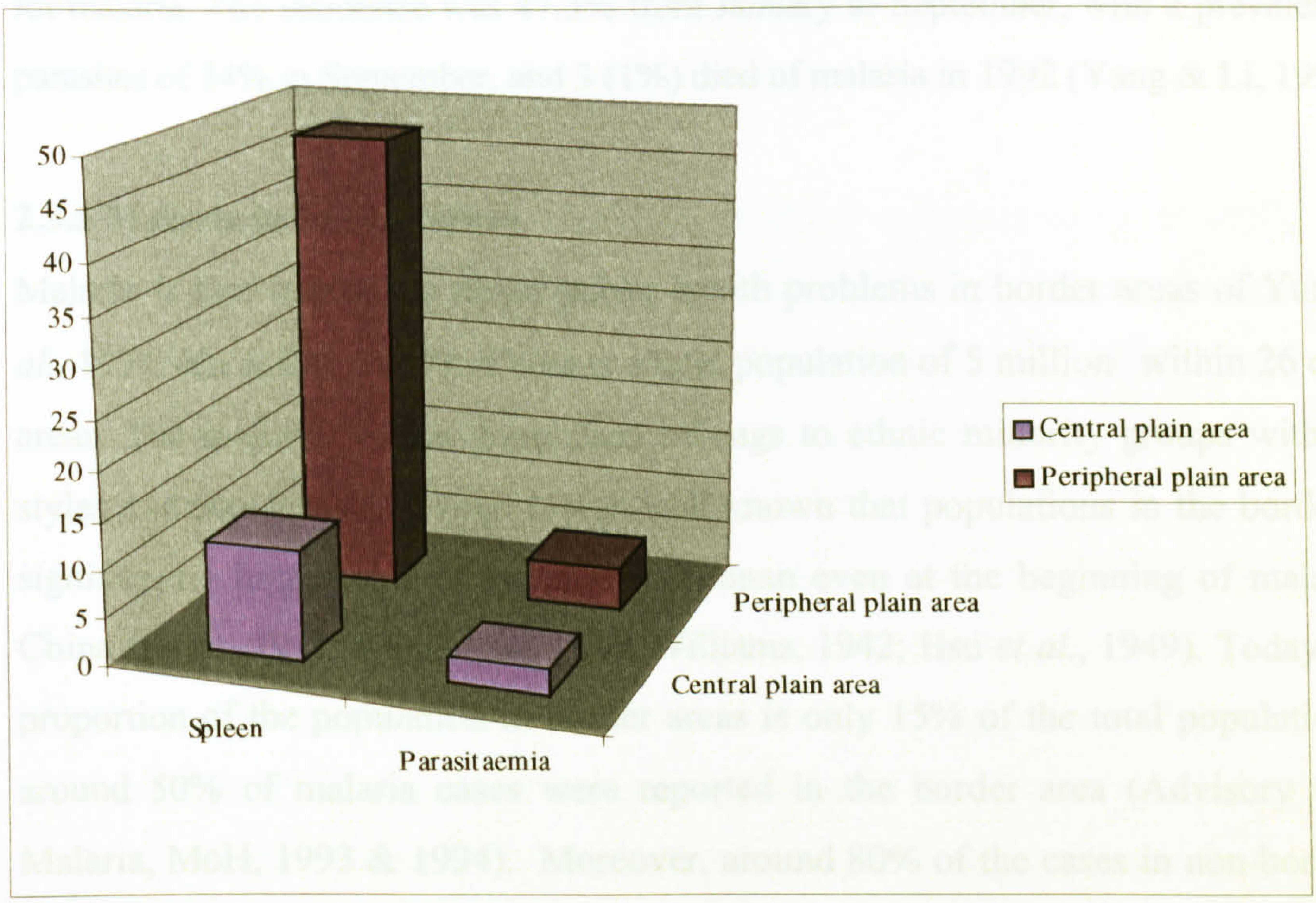
	Parasitaemia rate(%)		
County	*Plain*	*Valley*	*Mountain*
Xingpin	7.5	5.1	0.0
Jingu	14.6	5.7	0.3
Monglian	25.3	19.5	2.1





Sources: Zhou(1991)

Figure 2.9. Malaria variations in the different ecological features in Yunnan (1956-1959)



Source: Zhang (1990)

2.3.4. Forest malaria, the effect of deforestation, and of reforestation on malaria

It is well known that deforestation tends to change the level of malaria transmission (Walsh *et al.*, 1993). With a tropical climate in the Southern and Southeast parts of Yunnan, the government has been encouraging exploitation of the Torrid Zone for economic development in recent decades. A large number of people either from in the province or from other parts of China have been entering the areas in the hope of improving their economic status. However, most areas are economically underdeveloped with poor transportation systems, generally lack of primary health care and health facilities, and poor knowledge of malaria and its control. The populations mostly come from areas not endemic for malaria so they have little immunity and suffer a high degree of malaria morbidity and mortality. They come to the marginal areas and work either for small enterprises or nationally owned farms to deforest the area and plant bananas, or rubber trees. Coffee, tea, sugarcane and various tropical fruits are grown in plantations. In 1990, among 507 workers in Jiabutuo rubber plantation in Mangla county, the malaria incidence was 74.8% per year and the prevalence of malaria parasitaemia was 60.8%, *P. falciparum* comprised 43.[illegible]% of infection and the mortality was 1% (Che *et al.*, 1992). A survey in a sugarcane plantation in Mongla county showed the malaria incidence was 31.5% from May to November 1991. The prevalence of malaria parasitaemia was 36.4%, the plantation in 1992, 322 workers were all from areas that were malaria free or less endemic areas for malaria. The incidence was 47.3% from January to September, with a prevalence of malaria parasites of 24% in September, and 3 (1%) died of malaria in 1992 (Yang & Li, 1992)

2.3.5 Malaria in border areas

Malaria is also one of the major public health problems in border areas of Yunnan (Zhou *et al.*, 1994; Xu & Liu, 1997). There is a total population of 5 million within 26 counties in the areas. The majority of the population belongs to ethnic minority groups with different life styles and standards of living. It was well known that populations in the border areas had a significantly higher risk of malaria in Yunnan even at the beginning of malaria studies in China (Faust, 1926; Ling *et al.*, 1936; Williams, 1942; Hsu *et al.*, 1949). Today, although the proportion of the population in border areas is only 15% of the total population in Yunnan, around 50% of malaria cases were reported in the border area (Advisory Committee on Malaria, MoH, 1993 & 1994). Moreover, around 80% of the cases in non-border areas were imported from either the border areas or the Red River basin area in the province (Zhu *et al.*, 1994). Plains are the main endemic areas, while the incidence of malaria in the mountain and valley areas is relatively lower near the border.

One of key reasons why malaria has remained significantly higher in border areas is probably that the environmental, meteorological and ecological situations are in favour of malaria transmission. All the areas are located in the tropical zone with a significant amount of rainfall and much humidity, and malaria is perennially transmitted.

Huge migrations of people acted as a driving force for malaria transmission in border areas (Zhu *et al.*, 1994). It is estimated that some 10 million people are moving around in the border, many of who are non-immune to malaria (Xu & Liu, 1997). Many people cross the border into neighbouring Myanmar, P.R. Laos or Vietnam, either to work in agriculture, mining or the lumber industry, usually working in the forest areas with few malaria control measures, or to do business in neighbouring countries. When they come back to China, they bring malaria into local areas. A survey of a group of farmers engaged in farming in Myanmar in 1992 showed that the prevalence of malaria parasitaemia was 32.1%. Of those, 73.8% were due to *P. falciparum*. The farmers went back to China due to the malaria outbreak and several people died of malaria (Zhang *et al.*, 1992). A further survey indicated that malaria incidence was directly proportional to the length of stay in people going to Myanmar (Li *et al.*, 1987). The average parasite rate in febrile cases returning from foreign countries was 6.67% in 1995 (Xu & Liu, 1997).

Refugees are another reason for malaria instability in border areas. In 1987, 7838 refugees from Myanmar poured into the border area of China. The prevalence of malaria parasitaemia was 33.3% among the refugees, of which 81.6% was due to *P. falciparum*. Subsequently, malaria led to outbreaks in local inhabitants (Zhu *et al.*, 1994).

In summary, complicated landscape and meteorological characteristics make for a great variation of malaria spatial distribution in Yunnan in general. Human behaviour, however, also plays a very important role in the malaria distribution. A landscape and environment based model would help us to understand the malaria transmission and spatial distribution pattern, and thereafter, the model would help us to set up more effective spatial predicting systems in the province.

Chapter 3

Geographical information systems and remote sensing and their role in epidemiology and malaria control and research

3.1 Geographical information systems and their roles in epidemiology

Geographical information systems are computer hardware and software packages concerned with the capture, storage, manipulation, query, display and analysis of any and all types of geographical information (Aronoff, 1989; DeMers, 1997; Burrough, 1998). A typical GIS database consists of spatial features in space, a database that contains descriptive information about the features, and the internal and external topological connections of the features recorded. The most defining characteristic of a GIS is its strict link between a feature's geographical position (co-ordinate) and its attribute data. This allows its users the ability to manipulate, analyse and query information on various map features through their shared components. The power of GIS is its ability to integrate and manipulate multiple layers or themes of spatial data for a large area. The data need not be from the same source, same map scales or same projects. The data that could be integrated into a GIS include paper maps, aerial photographs, satellite images, GPS data as well as routine data, census, epidemiological surveys, environmental data, health and many data from other studies. In all cases the data must be registered geographically to a generally accepted and coded common frame of reference (usually a geodetic co-ordinate system), before they can be stored in a GIS. By facilitating the manipulation and possible combination of several data layers from the disparate data sources, a GIS permits rapid display and analysis of multivariate spatial data (Nicholson & Mather, 1996). The wide-ranging functionality of GIS clearly makes it a potentially powerful tool for epidemiology, particularly in the areas of, (i) mapping diseases and their determinants, (ii) quantifying risk variables, (iii) linking diseases and their potential risk variables, and (iv) creating databases for further statistical and epidemiological analyses.

3.1.1 Mapping diseases and their determinants

Mapping spatial and temporal distributions of diseases and their potential determinants is probably one of the most immediate uses of GIS. Historically, mapping diseases has provided important clues to the aetiology of diseases. Snow (1849), for example, developed his hypothesis on the mode of transmission of cholera by mapping the association between cholera mortality and the different water supply companies in London, as well as the association with London's Broad Street water pump. Burkitt (1962) published geographical distributions of cases of malignant lymphoma in Africa that helped to develop the hypothesis that an insect vector was important in its aetiology. The most commonly used maps in epidemiological study are point maps and choropleth maps (Cliff & Haggett, 1988). Individual data may be present as points in a GIS to form a point map. A choropleth map may be produced by allocating each case to its relevant administrative boundary. Maps provide instant impressions of the relationships of diseases (incidence or mortality) to their appropriate geographical positions (often as the surrogate of disease determinants). They can reveal spatial variations and distribution patterns not previously discerned nor suspected from the examination of tables of statistics (Howe, 1986). They provide essential descriptive information relating to geographical variations in disease incidence or mortality. They answer the epidemiological question 'where'?

A traditional map is static, and creating a map needs a skilled cartographer. Information is encoded in the form of symbols as points, lines or areas in a maps and the symbolism of colours or text codes is usually explained in a legend. It is created by a survey at a given point in time, and it represents a snapshot of natural environments. Once data had been put into a map, it usually had to serve for a long period of time: it was neither cheap nor easy to update such maps based on current changes. It is also difficult to retrieve the data from them in order to combine them with other spatial data or to create new information from existing maps. Besides, the original mapping data usually have to be greatly reduced in volume, or classified to make them understandable and representable. Consequently many local details were often filtered away through map generation (Burrough & McDonnell, 1998). Finally, because a printed map represents qualitative documents, it can be extremely difficult to attempt quantitative spatial analysis within the units delineated on a thematic map. GIS is more than a simple automator of the tasks of a cartographer of conventional maps. It holds a database in digital form instead of a conventional map printed on a piece of paper or film. Once the data have been stored in a database, they can be retrieved and transformed to provide information based on need. New data may be created from old and data can be used repeatedly without any deterioration in quality. Therefore, a GIS spatial database need not remain static, but may be used to model changes in spatial and temporal distributions of diseases and their determinants.

As the examples above suggest, a major aim in the study of geographical variation in disease incidence or mortality and its potential determinants is to formulate hypotheses about the aetiology of a disease by taking into account spatial variation in environmental factors. Disease maps can be readily converted into areas of risk. However, geographical association will never prove or establish causation (Howe, 1986). Mapped variation might be due to the effects of confounding variables or other biases such as under-reporting and misdiagnosis particularly, when the disease maps are based on routine data (English, 1992). Hypotheses developed in geographical analysis need to be tested using more rigorous studies.

Mapping diseases and their potential determinants might simply map the raw data to show where the diseases and their potential determinants are (Hu *et al.*, 1998; Cumming, 1999; Kidson *et al.*, 1999; Coetzee *et al.*, 2000), what are the patterns of the diseases and their determinants, what aspect of the disease has been changed and their determinant patterns since then? Using the interpolation technique of GIS is another common methodology to generate the disease pattern and their determinant surfaces (Cliff & Haggett, 1988, Nicholson & Mather, 1996; Estrada-Pena,1998), which will be discussed in detail in section of 3.1.2 in this Chapter. Recently, mathematical modeling in GIS has been widely used for mapping malaria and its vectors (Craig *et al.*, 1999; Snow *et al.*, 1999; Lindsay & Martens, 1998, Lindsay *et al.*, 1998), the leishmaniases and their vectors (Cross *et al.*, 1996; Thomson *et al.*, 1999), schistosomasis (Malone *et al.*, 1994; Malone *et al.*, 1996), dengue (Patz *et al.*, 1998), African trypansomiasis (Rogers & Randolph, 1993), and lymphatic filariasis (Michael & Bundy, 1997; Lindsay & Thomas, 2000). More will be discussed in the section of 3.3 in this Chapter

3.1.2 Quantitative risk variables

Typically, epidemiologists collect data for exposure at certain points in a study area or obtain exposure data from routine monitoring sites. GIS provides a range of spatial interpolation techniques to help epidemiologists to quantify exposure variables at unsampled or unmonitored sites, particularly for a large area through modelling point data, such as Thiessen polygons, linear interpolation, inverse distance-weighting interpolation, trend surface analysis, regression and kriging . (Bonham-Carter, 1994; DeMers, 1997; Burrough and McDonnell, 1998).

Thiessen (or Voronoi) polygons are a proximal interpolation technique that can be carried out as a the standard component of most proprietary GIS and perhaps one of the simplest and widely used interpolation techniques for estimating unsampled exposures in a GIS (Briggs, 1992). The basic characteristics of the polygons are that every location within them is closer to the point encircled than any other point in the coverage. Therefore, Thiessen polygons, in effect, are used to predict the values at surrounding points from single point observation (Aronoff, 1989). We could generate Thiessen polygons around point data in which exposure data are collected. The variables in the polygons could be estimated by the point exposure. However, the method has a number of limitations (Burrough, 1986; Aronoff, 1989; Briggs, 1992). The division of a region into Thiessen polygons is completely dependent on the location of the observation point. The value assigned to each polygon is estimated from a single sample. Estimates of error cannot be calculated from a single sample. Finally, Thiessen polygons estimate a value at unsampled points, but do not assume that points close together are more similar than that points far apart, an assumption that is usually appropriate in geographical analysis. However, Thiessen polygons are probably best for qualitative variables for which other interpolation methods are inapplicable or where the measurable local spatial variability is low (Burrough, 1986; Briggs, 1992).

In contrast to the discrete technique described above, almost all other methods of interpolation embody a model of continuous spatial change that can be described by a smooth, mathematically defined surface. The simplest way to describe gradual long-range variations (trend surface) is to model them by polynomial regression (Cliff & Haggett, 1988). To develop the trend, the whole coverage will be divided into regular grids and each of the values of the sampled exposure variable in the region is examined and fitted into a mathematical equation. From the equation used to estimate the surface of best fit, a single value is estimated and assigned to the regular grid in the coverage. The process continues for other targets, and the trend surface can then be extended for the entire coverage. The technology has been widely used in meteorological science to generate temperature and rainfall surfaces based on the point data from meteorological observatories. In epidemiology, the most obvious such features are probably air pollution, e.g. concentration of SO_2 in the vicinity of electricity generating plants. The very best which can be achieved is a set of pollution measurements at some sample locations. Within a GIS, a sample interpolation procedure can be use to generate the pollution 'surface'. The advantage of trend surface analysis is that it is a technique that is quite easy to understand. It reflects the 'mean' trend of modelled variables, embodies ideas of gradual change, and is a 'smoothing function'. However, for altitude data in rough terrain, such a technique would tend to smooth abrupt altitude change. It is mostly suitable for environmental variables not changing abruptly within a study region, such as aerial pollution. Similar methods are also available in the GIS environment, e.g. linear interpolation, inverse distance-weight interpolation and splines. The methods of interpolation described above present several limitations (Carrat & Valleron, 1992; Burrough & McDonnell, 1998). First, they fix tuning constants or make prior assumptions that do not take advantage of the spatial structure of the variable. Second, they do not allow for estimate of the error of interpolation. Thirdly, there is no *a priori* method of knowing whether the best values have been chosen for the weighting parameters or if the size of the search neighbourhood is appropriate.

Kriging, a technique widely used in earth science, provides a means of interpolating values for unsampled locations and calculating a measure of variance around estimated values. Kriging uses the idea of a regionalized variable, which varies from place to place with apparent continuity but cannot be modelled with a single smooth mathematical equation (DeMers, 1997; Burrough & McDonnell, 1998). It is a local interpolation technique, which is very sensitive to small-scale variation. The interpolation proceeds by first exploring and then modelling the stochastic aspects of the regionalized variable. The resultant information is then used to estimate the weights for interpolation (Burrough, 1986; Carrat & Valleron, 1992). Nicholson and Mather (1996) mapped nymphal ticks by using kriging to study the relationship between exposure to deer ticks and the risk of contracting Lyme disease in Rhode Island. They identified 80 sites located throughout the state that have been sampled annually, beginning in 1993, to estimate distribution and abundance, and *B. burgdorferi* infection rates of ticks. Data for each site were entered into the GIS for analysis. In this study, each point was interpolated by using the Gaussian model to weight eight surrounding sample points based on their distance from the interpolated point. The result of kriging modelling was a map depicting a continuously varying surface of nymphal tick density in the forested habitats of Rhode Island.

Most often, the point data are not only the source of information about the distribution of environmental variables to be interpolated and other extra information might help the interpolation with kriging techniques. Therefore stratified kriging, co-kriging and multivariate kriging have come into fashion (Burrough and McDonnel, 1998). An advantage of kriging is that it follows the basic principle of geographical science: points that lie close together in space are more likely to have similar values of a property of interest than points further apart (Burrough, 1986; Burrough & McDonnell, 1998). Kriging also provides an estimate of the potential amount of error of the estimates as well as the estimated values themselves. One important limitation of kriging is that it doesnot model variation at the scale smaller than the lag interval, often determined by the average spacing between neighbouring samples (Burrough, 1986). Thus, important local variation can be masked by this technique.

3.1.3 Linking exposure and disease, creating a database

GIS offers a variety of options for spatial analysis including the ability to overlay several map layers, buffering, distance and area calculation When multiple geographical data are stored in a common coordinate system, many map layers can be merged and viewed simultaneously, and both spatial data and their attribute data can be joined together through the overlay process. This allows the user to look through the set of maps and better understand the spatial relationships among the features of the different layers. From an epidemiological study point of view, this approach can be used visually compare maps of disease with those of exposure variables, creating a database linking exposure variables with diseases. Cliff and Haggett (1996), for example, demonstrated this by re-evaluating John Snow's classical cholera data in London in the 1880's within a GIS. The GIS was used to develop two maps. The areas of the city situated 25 feet or less above the River Thames they expected to be able to relate to the level of cholera death due to proximity to the River Thames (a defective-drainage factor causing sewage to be discharged ineffectually from the houses). The second map classified areas which areas had water supplies provided by each of the water companies (the water companies drew water from different part of the River Thames). They overlaid the two maps and calculated mortality for each area. A substantial excess of death was seen in those areas covered by both factors. If a single factor is considered, then it appears that defective drainage is more important than is the Water Company supplying the district. Areas in which neither factor operates have a substantial deficit of death. Recently, Glass *et al.* (1995) used a GIS to overlay six different environmental data, including land use/land cover, forest distribution, soil, altitude, geology and watersheds, creating a database containing 53 environmental variables in order to identify and locate residential environmental risk factors for Lyme disease in Baltimore county, Maryland. They also entered the address of 48 cases and 495 randomly selected control addresses into the GIS. Residential information for cases and controls was obtained from the database created from the GIS overlay. The variables: residence in forested area, on Glenville/Gleneg soils, and certain specific places, were associated with elevated risk of getting Lyme disease. Residence in highly developed regions was protective.

Buffering is the process of creating areas of calculated distance from a point, line, or area object and is a widely used GIS function to link exposure with disease. The use of buffering techniques can simply divide a large group of population into 'exposed' and 'non-exposed' groups. Morbidity or mortality of the disease of concern may then be compared between groups. In other cases, risk may be assumed to vary with distance from point pollution sources, or distance from vector breeding places. Therefore, tests are required to determine whether the distribution of diseases is inversely related to distance away from pollution sites and vector breeding places. Dunn *et al.* (1995) used buffering to study the association between prevalence of asthma and factory pollution in a small town of County Durham. They found that the area within 1km and immediately the north east of the factory had a significantly higher age and sex standardised asthma prevalence than expected. The evidence clearly supported the hypothesis that the factory pollution might contribute to the excess prevalence of asthma in the study area. le Sueur *et al.* (1997) used GIS to study the relationship between malaria incidence and irrigation. They used GPS to obtain the geographical locations of homesteads in Mamfene, South Africa. The location of each household was input into the GIS and the number of cases per homestead for 4 months in 1993 were also input into the GIS. A 2km buffer was generated around the suspected exacerbating feature, resulting from the irrigation scheme, to estimate vector mosquito range. Their results showed that almost all the malaria cases fall within this buffer area, thus supporting the hypothesis of association between the scheme and malaria incidence. Thompson *et al.* (1997) used buffering techniques to indicate that the risk of malaria was 6.2 times higher for an individual living less than 200 metres from breeding sites than for an individual living 500 metres or more away from the breeding sites in a suburban area of Maputo, Mozambique.

Traditionally, geographical analysis is an essentially ecological study. Aggregated disease data are correlated with aggregated environmental data. Most estimates of the distribution of mortality and morbidity are based on the assumption of homogeneous distribution within large administrative units. In current epidemiological methodologies there are relatively few techniques for using or synthesising the discriminating data from small geographical units (villages or individual household) in a statistically acceptable manner (Elliott *et al.*, 1992). GIS techniques make the estimation of individual environmental exposures possible. Glass *et al.* (1994) evaluated association of the environmental exposure and the abundance of *Ixodes scapularis* on deer. They used Thiessen polygons to assign the deer's environmental exposure. Glass *et al.*(1995) used an overlay technique to identify each Lyme disease case and their four controls exposed to environmental variables. Pikhart *et al.* (1997) studied the association between ambient air concentrations of nitrogen dioxide and respiratory symptoms of school children in the Czech Republic. He assigned the individual child's exposure to nitrogen dioxide by using buffering around school and home.

In summary, GIS is a powerful tool which can help epidemiologists and public health experts on disease control to visually assess the spatial aspects of public health problems. It is also good for communicating these issues to decision makers. Mapping the features of disease in GIS will help epidemiologist to answer: where is the problem and what spatial patterns exist? What has changed over time? Besides, GIS can help epidemiologists to quantify the environmental variables by a set of interpolation techniques and spatial analysis in the GIS environment. The spatial analysis of GIS, such as overlaying, buffering and distance calculation can help epidemiologists to link the diseases with their environmental determinants and subject them to further statistical analysis and mathematical modelling within and outside of a GIS environment. GIS modelling will help epidemiologists to forecast the trend of the disease spatially and temporally. They answer the epidemiological questions: what if circumstances change, and what will be the consequences of environmental variation?

3.2 Remote sensing and mosquito population dynamics

Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device (sensor) that is not in contact with the object, area, or phenomenon under investigation (Lillesand & Kiefer, 1994). The sensors are mostly cameras or satellite-borne multi-spectral scanners or radar. The sensors are mounted on a "platform" from a few metres high, through aircraft, rocket and space shuttle thousands of metres up, or located on satellites hundreds or thousands of kilometres above the subjects of interest. Most notable are aircraft, earth-orbiting satellites such as Landsat and the SPOT series, and polar-orbiting, meteorological satellites such as the National Oceanographic and Atmospheric Administration, Advanced Very-High-Resolution Radiometers (NOAA-AVHRR) and Meteosat, High Resolution Radiometer (HRR). The science of remote sensing rests on the fact that every object in nature reflects, absorbs and emits energy at specific and distinctive wavelengths in the electromagnetic spectrum. Passive sensors, including cameras and multi-spectral scanners, record reflected or emitted energy and the active sensors such as radar transmit microwave energy to the subject, and then record the return signal. Remote sensing provides the capability to collect uniform measurements in digital form over large areas at very high speed and to analyse phenomena that could not be monitored in any other way (Aronoff, 1989). Although remote sensing had been argued to be a useful tool for epidemiological study of a disease long ago, particularly for vector-borne diseases (Cline, 1970), the technology was only preliminarily used in the study of vectors and vector-borne diseases as extensively reviewed (Hugh-John, 1989; Washino & Wood, 1994; Roberts & Rodriguez, 1994; Hay *et al.*, 1996; Hay *et al.*, 1997; Thomson *et al.*, 1996 & Thomson *et al.*, 1997). The present review focuses on its roles in studying mosquitoes, particularly in (1) mapping mosquito habitat, (2) and the study of mosquito abundance and population dynamics.

3.2.1 Mapping mosquito habitats

Most early studies using remote sensing to study mosquitoes were to identify and map their habitats based on landscape, vegetation, water and soil characteristics, which had already been documented to be mosquito breeding sites. One of the earliest studies done by National Aeronautics and Space Agency (NASA) scientists used colour-infrared (CIR) aerial photography to identify *Aedes sollicitans* breeding habitats near New Orleans, Louisiana in 1971 (cited by Hugh-Johns, 1989). This study documented a strong empirical relationship between *Bacopa monnieri* (water hyssop) and the mosquito. Wagner *et al.* (1979) later used CIR aerial photography to map forested wetlands, open marshes and residential areas in parts of Michigan. By combining the maps with information on population density and flight range of *Ae. triseriatus* (the vector of California encephalitis) and *Culex* spp. (vectors for St. Louis encephalitis), the maps were then used to direct mosquito control measures. Welsh (1989) also used CIR aerial photography as a survey technique for egg populations of *Psorophora columbiae* in rice fields of Louisiana and Texas. The resultant photographic data clearly indicated that landscape and terrain features such as rice field levees, tyre tracks and ditches were attractive for oviposition by *Ps. columbiae*. More recently, Dale and Morris (1996) used similar methods for rapidly mapping the breeding sites of *Cx. annulirostris* (the vector of Ross River virus) in urban areas of Queensland. The accuracy and completeness of identification were 87% and 75% respectively, with reference to the original field site identification and by further field checks.

Barnes and Cibula (1979) extended the NASA study to include the use of automated computer processing techniques for analyses of digital multi-spectral scanner (MSS) data acquired from aircraft. Using these data and techniques, they were able to map favourable egg-laying habitats for *Ae. sollicitans* in the Mississippi Delta. Hayes *et al.* (1985) first used Landsat 1 and 2 MSS data to identify *Culex tarsalis* and *Ae.vexans* habitats at Lewis and Clark Lake in Nebraska and South Dakota. They found the mosquito larval habitats associated with fresh-water plant communities, wetlands, and other aquatic features in the study areas. The resultant scenes, by a supervised classification, identified these aquatic habitats with 95% accuracy.

More intensive studies have been carried out to map Rift Valley Virus vector mosquitoes in Kenya (Linthicum *et al.*, 1987; Linthicum *et al.*, 1990; Linthicum *et al.*, 1991; Pope *et al.*, 1992). Preliminary analysis revealed that NDVI (derived from NOAA-AVHRR) were positively correlated with seasonal rainfall pattern, and such rainfall can flood mosquito breeding habitats, known in Kenya as "dambos" (Linthicum *et al.*, 1987). These habitats were highly suitable for *Aedes* and other mosquito breeding. A further study showed that NDVI values at or above 0.43 corresponded to at least short-term flooding of vector mosquitoes' breeding habitats (Linthicum *et al.*, 1990), and this was used as a basis for mapping mosquito vector habitat flooding over the whole of Kenya. The study was extended using higher spatial resolution satellite imagery data (Landsat and SPOT), to map individual areas of mosquito breeding sites (Linthicum *et al.*, 1991; Pope *et al.*, 1992). Unsurprisinly, dambos could be identified with greater accuracy using the products, although there was still no way to differentiate between unflooded and flooded dambos. For this, the authors used airborne Synthetic Aperture Radar (SAR), tested on artificially flooded sites.

Ritchie *et al.* (1993) used land-based radar to predict *Ae. taeniorhynchus* habitat flooding status in Collier, Florida. They used radar data to estimate rainfall, and combined rainfall estimates with tide gauge data, to serve as criteria for the sites for larval inspection. Rain, rain plus tide, and tide triggered 48, 26 and 26% of the 14 proposed inspection trips. Seven out eight larval breeding sites were detected by the system.

3.2.2 Mosquito abundance and population dynamics

Later remote sensing studies have tended to concentrate on using remote sensing data to identify and monitor environmental factors influencing vector populations. Most of the studies integrated remote sensing and ground survey data into GIS, and studied the spatial and temporal variability in population dynamics of mosquitoes, thereby predicting mosquito population abundance and dynamics and estimating the potential for disease transmission. NASA scientists were the first to extend the study from mapping mosquito habitats to their environmental determinants. In 1985, a pilot study was carried out by NASA scientists in California's Lower Sacramento Valley rice fields (Wood *et al.*, 1991). Digital airborne Thematic Mapper Simulator (TMS) data of the study area were acquired throughout the growing season. The larval densities of *An. freeborni* were obtained in 46 rice fields throughout the study period. Normalised difference vegetation index (NDVI) was first used to study malaria mosquitoes in this study. The initial results indicated that rice fields with high larval populations were often associated with early development stages of the vegetation canopy. The distinction could be made over two months before peak mosquito production. In addition, over 70% of high mosquito-producing fields were located within 1.5km of a livestock pasture where adult mosquitoes could find blood meal sources. The subsequent work at a large scale was done in the same area in 1987 (Wood *et al.*, 1992). They revealed that the spectral reflectance data from Landsat 5 TM in early-season canopy development and distances between rice fields and livestock from GIS could be used to distinguish between the high and low mosquito-producing rice fields. Using spectral and distance measures in either discriminant or Bayesian analysis, the identification of high mosquito-producing rice fields was made with 85% accuracy nearly two months before the mosquito larvae peaked.

More studies were carried out aiming to relate mosquito abundance to its environmental determinants in Mexico. Beck *et al.*(1994) used GIS combined with Landsat TM data to discriminate between villages at high and low *An. albimanus* abundance in an area in southern Chiapas, Mexico. Land use and cover maps were generated by unsupervised classification of Landsat TM data. The buffering function of GIS was used to determine the proportion of mapped landscape elements surrounding 40 villages where *An. albimanus* abundance data had been collected. Multivariate analyses indicated that transitional swamp and unmanaged pastures were two of the most important landscape elements determining vector abundance. The discriminant functions generated for the two elements were able to correctly distinguish between villages with high and low mosquito abundance, with an overall accuracy of 90%. The models were verified in a neighbouring area (Northwest of Huixtla), two years later (Beck *et al.*, 1997). Although the remotely sensed landscape composition indicated some differences in terms of the relative proportions of landscape elements, the discriminant model accurately predicted 79% of high abundance villages and 50% of the low-abundance villages for an overall accuracy of 70%. The regression model correctly identified seven of the ten villages with highest mosquito abundance. Rodriguez *et al.* (1996) carried out a similar study in the coastal plain areas of Southern Chiapas, Mexico. They studied the relationship between landscape characteristics around human population centres and adult *An. albimanus* abundance in 14 villages sampled at weekly intervals. A land use and land cover map was created on the basis of the aerial photographs from NASA ER-2 aircraft. One kilometre buffers for all 14 villages were created in a GIS. Multivariate analysis (using a tree-based model) indicated that flooding in unmanaged pastures is determined by their altitudes as well as that the presence of transitional swamps and mangroves was the most useful landscape indicator of *An. albimanus* abundance.

Rejmankova *et al.* (1995) used SPOT data to map landscape features in a study of adult *An. albimanus* densities in northern Belize. Unsupervised classification was employed to classify the SPOT image data. Results indicated that marshes sparsely populated with macrophytes have dense cyanobacterial mats and were very productive for *An. albimanus* larval habitats, and that the river margin was also an important mosquito habitat. Based on the distance between human settlements and the nearest marsh/river exhibiting this particular class combination in the satellite image, they predicted sites less than 500 m from habitats as having a high mosquito population density, and sites over 1,500 m distant having a lower density. The predictions were verified by collections of mosquitoes landing on people. The prediction was 100% accurate in the lower category and 89% accurate for sites in the higher category. Roberts *et al.* (1996) carried out a similar study for predicting the spatial distribution of *An. pseudopunctipennis* in 14 villages of central Belize. Based on environmental criteria, distance of houses from waterway, altitude above specified waterways, and amount of forest between houses and waterways, all of which could be identified either on satellite images or from existing maps, four of eight sites that were predicted to be high probability locations for presence of the mosquito were positive and all low probability sites were negative.

Sharma *et al.,* (1996a; 1996b & 1997) used the Indian Remote Sensing satellite to study the mosquito distribution around Delhi, India. The water bodies with marshy areas, vegetation and human settlement were considered to be primarily responsible for mosquito abundance. Supervised classification was used to produce land use and land cover maps. A survey of larval and adult mosquito density was performed concurrently in the study sites. Results indicated that the spatial variation of mosquito densities was positively correlated with water bodies and vegetation in some study sites.

Gleiser *et al.* (1997) recently used NDVI generated from NOAA-AVHRR and other data to study the abundance and dynamics of *Ae. albifasciatus* (the vector of western equine encephalitis) in the south of Mar Chiquita Lake of Argentina. The adults, larvae and pupal stages of the mosquito were sampled at 5 sites. The results of correlation analysis between adult and pre-adult mosquito densities and meteorological data and NDVI indicated that NDVI was significantly correlated with the mosquito abundance, and particularly the pre-adult stages. As NDVI could be used to estimate larval abundance, it could also be used to predict adult abundance seven days in advance.

Thomson *et al.* (1997) studied the spatial relationship between 2*R* chromosomal inversion polymorphisms of *An. gambiae* using NDVI and rainfall in 16 sites throughout Mali. They used monthly composite NDVI values obtained from each collecting site at the time of sample collection. The analysis indicated a highly significant correlation between the *2R* inversion in the population and NDVI values, with higher frequencies being associated with low NDVI value.

In conclusion, the use of remote sensing data in mapping mosquito spatial and temporal distribution and abundance has great potential. But most of the present studies are on a very small scale and lack validating studies. Nevertheless, a system that integrated remote sensing data with other data sources such as meteorological and landscape data, into GIS would be a potentially powerful tool for mosquito mapping, and assessing abundance and dynamics in the future.

3.3 GIS, remote sensing, malaria mapping, stratification, surveillance and prediction

Recent re-assessment of the global malaria control strategy has led to the development of more flexible and more reliable approaches to malaria control, which consider the requirement of local conditions (WHO, 1993). An essential element in this development is recognition of variability of environmental and epidemiological parameters, which influence the pattern of malaria transmission risk. One of the four key technical elements of the strategy is to detect epidemics early so as to contain or prevent them. Therefore, new tools are urgently needed to help health planners to study the variability of malaria, set up effective surveillance systems, and predict malaria epidemics. GIS and remote sensing may assist in this process.

3.3.1 GIS, remote sensing and malaria mapping

Mapping of malaria has long been considering an important step for malaria control programs. A simple sketch map of malaria could highlight where high risk malaria areas are and where malaria tends to be epidemic, so that we need to keep alert, and where malaria control should be focused and where is no malaria transmission. The integration power of GIS and the digital nature of its database make it particularly suitable for mapping and easily updating the risk of malaria and to guide malaria control services. One might simply map malaria distribution with the available routine data. The simple map will help public health authorities to know where malaria is, what are the patterns of malaria distribution and how high the risk of malaria is in different places, as well as recording any changes in malaria spatial distribution patterns. This simple information will help malaria control authorities and malariologists to identify where malaria control should be focused to achieve maximum benefit. Due to the unreliability of routine data (mostly underreporting), we might use point data to model malaria distribution based on interpolation technology in a GIS as discussed in section 3.1.1. We could also map malaria transmission and spatial distribution based on its determinants by mathematical or statistical modelling by using variables such as altitude, or land use, or temperature, or rainfall or satellite imagery derived information . Recently, the mapping malaria risk in Africa (MARA) initiative has been using GIS to model and map the malaria risk in the whole continent of Africa (Snow *et al.*, 1996; le Sueur *et al.*, 1997). They used fuzzy logic climate suitability to model the malaria transmission in Africa, based on biological constraints of temperature and rainfall on malaria parasites and their vectors' development (Craig *et al.*, 1999). The continental malaria suitability map derived from the fuzzy logic climate suitability model compared well with contemporary field data and historical 'expert opinion' maps. They combined a review of the literatures on malaria surveys in Africa and the resulting continental risk map derived from the fuzzy logic model to estimate malaria morbidity and mortality in Africa (Omumbo *et al.*, 1998; Snow *et al.*, 1999a; Snow *et al.*, 1999b). Although the model is still crude with low resolution (5 x 5 km^2), . it might go beyond 'best guess'.

3.3.2 GIS, remote sensing and malaria stratification

Stratification is the process of uniting strata (area, population, situation) that have in common a set of specified characteristics (WHO, 1986; Singh *et al.*, 1990). The ideas of stratifying malaria originated from the features of epidemiological zones in terms of their determinants. Many factors may be stratified and contribute to the epidemiology of malaria, such as the distribution and prevalence of parasite species, anti-malarial sensitivity of parasites, the distribution, abundance and efficacy of vector species and their behavioural characteristics, the intensity of transmission, as well as ecological and geographical characteristics of different areas such as landscape, land-use patterns and land-cover, and meteorological conditions.

One of the main functions of GIS is the selection of areas based on some logical and mathematical rules according to some specific criteria. The criteria may be spatial or nonspatial. Queries may be developed to stratify an area according to the incidence of malaria and intensity of malaria transmission, the distribution and abundance of malaria vectors, topographical features, various land uses and land covers, hydrology, socio-economic situation, etc. Characteristics may also be combined. For example, we might stratify areas that have low altitude, have rice paddies that favour mosquito breeding and that are within 3km of a water reservoir. Each of these attributes comprises a distinct data layer. With GIS, one can create a 3km buffer around the reservoir and then select areas meeting all three criteria. Richard (1993) used AtlasGIS to stratify onchocerciasis in an endemic area of Guatemala. Based on the knowledge that temperatures below an altitudes of 500 metres are too hot to support high vector densities, while above 1,500 metres, cooler temperatures retard vector biting frequency and parasite development, he selected the communities lying between 500 and 1,500 metres by using the GIS. The simple GIS query function stratified 2,939 communities into higher and lower risk strata; only 1,288 communities were classified into higher risk communities in the endemic area.

Malaria stratification should be dynamic and ongoing, with the capacity to accommodate both expected and unexpected changes (WHO, 1986). In reality, however, malaria stratification tends to be static and is often used over periods of decades despite the fact that malaria and environmental and socio-economic situations have changed. This is understandable either due to the difficulty of monitoring such changes, particularly for environmental data over large areas, or the long time needed for compiling, analysing and presenting the data set. Remote sensing of continuing measurements of environmental and ecological variables on a relatively short temporal cycle makes it particularly valuable for monitoring environmental variation, although cost issues remain. The digital nature of the GIS database facilitates updating of malaria distribution to incorporate remotely sensed environmental data, ecological situations and other surveillance data. Therefore, GIS and remote sensing make 'ongoing' stratification possible and much easier than before.

3.3.3 GIS, remote sensing and malaria surveillance

Malaria surveillance involves continuous malaria case detection, parasitological examination, anti-malarial drug treatment, epidemiological investigation, entomological investigation, malaria control, and assessment of the effectiveness of malaria control measures (WHO, 1986). Malaria surveillance systems can provide timely support for decisions concerning control of endemic malaria, identification of malaria epidemics and evaluation of malaria control measures. Many applications of malaria surveillance systems involve the collection, compilation, and analysis of large quantities of data with spatial properties. GIS can provide a suitable framework for integration of spatial variables from various sources, including malaria cases, mosquito breeding sites, occurrence of antimalarial resistance and mosquito resistance to insecticides, demographic data, administrative boundaries, topography etc. A further advantage is that graphical displays are accessible to decision makers, helping them interpret the problem and decide on appropriate interventions. Remote sensing is particularly powerful in monitoring mosquito-breeding sites, predicting mosquito abundance and population dynamics, as reviewed in the section of 3.2 of this Chapter. However, a comprehensive spatial surveillance/stratification system can only be achieved if various data sources are integrated within a GIS.

Kitron *et al.* (1994) developed a national surveillance system for breeding sites of *Anopheles* mosquitoes and imported malaria cases by using GIS in Israel in 1992. The database in the GIS surveillance systems included coverage of all settlements and roads in Israel, administrative boundaries, mosquito breeding sites, and location of imported cases. Mosquito data were routinely collected from all potential *Anopheles* breeding sites in the country between March and November. The imported malaria data were entered into an existing database file of all population centres in Israel. Any population centre with at least one malaria case was considered positive. All data on mosquito breeding sites and location of imported malaria cases were entered into DBASE files that were read by MapInfo. They calculated the distance between the positive population centres and mosquito breeding sites, and assessed the risk of malaria transmission with consideration of vector capacity and flight range of each *Anopheles* species. The GIS-based surveillance systems ensure that if a localized outbreak were to occur, it could easily be associated with likely breeding sites, a specific *Anopheles* vector and probable human source, so that prompt control measures could be mostly efficiently targeted.

Nobre *et al.* (1997) developed their own GIS software, GISEip, for malaria surveillance by accurately identifying where and when malaria cases occurred within communities in the Amazon region, Brazil. More recently, Ghebreyesus *et al.*, (1999) used GIS to spatially analyze malaria distribution and monitor the coverage of community based control activity in relation to the population at risk in Tigray, Northern Ethiopia. They integrated population, health facility and community based malaria surveillance systems in the area into GIS

ArcView. This facilitated immediate visualization of the number of malaria cases in the community and helped to identify malaria 'hot spots'.

3.3.4 GIS, remote sensing, malaria spatial and temporal prediction

Point malaria data such as malaria parasitaemia, malaria incidence and mortality, are quite common in active malaria surveillance processes. GIS provides a range of interpolation techniques, as already discussed in section 3.1.2, which can use the point data to predict the malaria profiles in unsampled sites. The simplest method of interpolation is to use Thiessen polygons. The techniques assume that all important variations are homogeneous and isotropic within polygons (Cliff & Haggett, 1988). Other methods of interpolation embody a gradual spatial change method, such as linear interpolation, trend surface analysis, inverse distance weighted interpolation and kriging (Cliff & Haggett, 1988; Carrat & Valleron, 1992; Ribe iro *et al.*, 1996). Nevertheless, the predictive methods described above are purely mathematical modeling based on point data and its covariate distances among sampled points without consideration of biological mechanisms. Malaria prediction should be based more on its biological mechanisms, and these determinants can be modeled.

The principal factors that influence malaria spatial and temporal distribution are climate data such as temperature, rainfall and humidity (Dutta & Dutt, 1978; Onori & Grab, 1980; Molineaux, 1988; Bruce-Chwatt, 1991), which regulate the biology of the development of both mosquito and parasite. Traditionally, weather is usually measured at ground level for malaria spatial and temporal variation analysis and prediction studies (Gill, 1923; Swaroop, 1946; Bouma & van der Kaay, 1994; Bouma & van der Kaay, 1996). With the recent advances in remote sensing, a set of satellite surrogate climate variables can be generated from satellite images. Temperature, rainfall, humidity, cold cloud days (CCD), land surface temperature (LST), sea surface temperature (SST), sea surface height (SSH) and NDVI can be obtained from meteorological satellites such as NOAA-AVHRR and Meteosat-HRR (Hay *et al.*, 1996; Thomson *et al.*, 1996; Thomson *et al.*, 1997; Lobitz *et al.*, 2000). Greenness, wetness and brightness can be obtained from high resolution satellites such as Landsat and SPOT (Dister *et al.*, 1997; Mather, 1999). Hay & Lennon (1999) compared meteorological data of LST, atmospheric moisture and CCD derived from AVHRR of NOAA with estimated land temperature, atmospheric moisture and rainfall derived from spatial interpolation of the data from meteorological stations. They considered that there was no conclusive evidence for optimality of either technique. GIS allows modern malariologists to combine malaria data with routine ground weather data as well as meteorological satellite data to study malaria spatial distribution in the broad climatic zones or regions. However, the spatial and temporal variations of malaria in the local areas or regions are strongly influenced by altitude, land use, land cover, deforestation, and soil etc. (Fonaroff, 1982; Molineaux, 1988; Service, 1991; Walsh *et al.*, 1992; Bruce-Chwatt, 1991). High resolution imagery, such as from Landsat TM or SPOT, is particularly helpful for use of GIS to study malaria spatial distribution associated with smaller landscape features such as villages, rice fields, deforestation and reforestation by updating land use and land cover patterns and landscape features in a GIS.

It is expected that changes of environmental components will affect the scale, nature and trends of tropical diseases, especially for malaria (Bradley, 1993). Global warming will tend to increase transmission at higher altitudes on mountains in tropical areas and the spread of transmission towards temperate regions with low or no transmission (Bradley, 1993; McMichael, *et al.,* 1995; Lindsay & Birley, 1996; Patz, *et al.*, 1996). The El Niño Southern Oscillation (ENSO), a global climate anomaly event, has been postulated to increase risk of malaria epidemics in the Indian subcontinent, Venezuela and Colombia (Bouma & van der Kaay 1994; Bouma & van der Kaay, 1996; Bouma *et al.*, 1997). The local environmental changes, such as deforestation, land use and land cover changes, will affect local malaria transmission dynamics (Service, 1991; Bruce-Chwatt, 1991; Walsh *et al.*, 1992; Molyneux, 1997). GIS and remote sensing are particularly valuable to study and predict malaria in currently changing environmental situations both locally and globally. The combination of landscape features and a global warming model in a GIS might allow one to model and then predict malaria spatial and temporal distribution locally or globally. Low-resolution satellite images such as AVHRR-VOAA can be used to monitor global sea surface temperatures and then identify ENSO (Parkinson, 1997), and NDVI can be used to study interannual climate variability and map the spatial patterns of ENSO related events (Anyamba & Eastman, 1996). Therefore, GIS, combined with remote sensing data and other data sources, might give a better spatial and temporal malaria prediction under the effect of ENSO. Current deforestation, and land use and land cover pattern changes could be easily monitored by an earth-orbiting satellite such as Landsat or SPOT and the polar or meteorological satellites AVHRR-VOAA and Meteosat-HRR (Skole & Tucker, 1993; Lillesand & Kiffer, 1994). GIS, integrated with the remote sensing and other data, will allow us to study the effect of local environmental changes on malaria transmission dynamics in relatively large areas.

Sharma and Srivastava (1997) used the remote sensing imagery and GIS to study the receptivity and vulnerability of a village to malaria in Sanjay lake and surrounding areas in Delhi, India. Based on their GIS analysis, they developed location specific malaria control strategies in their study areas. Thomson et al (1996) used GIS IDRISI software to integrate NDVI, CCD from meteorological satellites NOAA-AVHRR and Meteosat-HRR respectively, with the alluvial soil map and irrigation system map as well as malaria epidemiological and entomological data in The Gambia. The preliminary results indicated that the values of NDVI could be used to identify possible mosquito breeding sites. Estimates of rainfall from CCD give a good approximation of temporal and spatial patterns of rainfall. The NDVI and CCD were highly correlated with the abundance of mosquitoes. The paper suggested that low-resolution satellite data integrated with other data sources could help explain spatial and temporal variation in malaria transmission patterns and thereby provide a basis for the study of malaria variability, malaria environmental stratification and early warning. More examples of using GIS and remote sensing to predict malaria vector abundance and population dynamics were given in section 3.2 of this Chapter.

Thomson *et al.* (1997) studied the temporal relationship between NDVI derived from NOAA-AVHRR and clinical cases of malaria in Niger. The results indicated that there were considerable degrees of seasonal association. They concluded that satellite data might provide a good indication of the timing and length of the malaria transmission period and form an early warning system.

Hay *et al.* (1998) used the NDVI, LST and CCD data derived from NOAA-AVHRR and Meteosat-HRR to model malaria seasonality in 5 community sites in Kenya. They compared the satellite derived variables with the mean percentage of total annual malaria admissions recorded in each month. They found that the NDVI in the preceding month correlated most significantly and consistently with malaria presentations across the 3 sites. They identified that a NDVI threshold of 0.35-0.40 was required for more than 5% of the annual malaria cases to be presented in a given month in the linear regression analyses. These thresholds were then extrapolated spatially with NDVI data to define the number of months, during which malaria admissions could be expected across Kenya in an average year after masking the non-transmission zone with DEM data.

In conclusion, GIS, integrated with remotely sensed and other environmental data, shows great potential for mapping malaria and its vector distribution, malaria risk stratification and surveillance. GIS and remote sensing could be useful tools for malaria spatial and temporal prediction through GIS and mathematical modeling. Nevertheless, malaria is a multi-risk factor disease. It is especially easily affected by socio-economic and human behavioral factors. Most previous studies focused on the landscape and environmental variables in their models although some started to incorporate some socio-economic variables in to their models (Thomson *et al.*, 1999). Future models should incorporate socio-economic and human behavior variables to a greater extent.

Chapter 4

Phase I Study

4.1 Objectives

1. To map the principal landscape and environmental characteristics in the study area.
2. To develop a geo-referenced database of malaria based on routine data.
3. To combine information from objectives 1 and 2 to assess the contribution of environmental factors to the variability of malaria in the study area.
4. To incorporate the results of the phase I study into specific hypotheses for the main field study (phase II).

4.2 Methodology

4.2.1 Study design

The phase I study is primarily an ecological study, and explores the correlation between landscape, environmental, ecological factors and the malaria spatial distribution. This phase involved visiting various relevant agencies, organisations, health and governmental authorities to access likely sources of data. The analysis is dependent on analysing routine malaria data and existing landscape and environmental data (maps). Data collection was carried out from May to June 1997

4.2.2 Study area and population

The phase I study was carried out in the county of Yuanyang in the Red River basin area, Yunnan (Figure 4.1). Yuanyang is situated between 22° 49'N and 23° 19' N, between 102° 27'E and 103° 19'E, and located in the middle and lower reaches on the western side of the Red River, approximately 150km north of Vietnam. The distance from South to North is 56 km and from East to West is around 54 km. The county has a total area of 2,200km^2. Its environmental and malaria profiles are broadly similar to those of the rest of the Red River basin area. This county was selected because of

a) Relatively high malaria endemicity.
b) Co-endemicity of *P. vivax* and *P.falciparum* in most of its area.
c) Terrain, land use patterns are representative of much of the Red River basin area.
d) Relatively reliable data set on malaria over past decade.
e) Availability of environmental maps, such as terrain maps, land use maps, soil maps etc.

Figure 4.1. The geographical location of Yuanyang County, Yunnan, China

4.2.2.1 Topography

Yuanyang County is a mountainous area with a vertical landscape and climate sequence, the highest altitude is 2,939 metres and the lowest 144 metres in the county. The terrain is steep, almost no flat area within the county. Generally speaking, the gradient of terrain of the county is over 20°. The Red River forms the boundary between Yuanyang county and Janshui county and Gejiu city in the north (Figure 4.1). It has a boundary with Jingping county in its south and southeast parts and with Honghe county to the west. There are various land-use types in the county, such as paddy fields, dry crop fields, and various tropical fruit plantations etc. The forest cover is around 25% of the total land area.

4.2.2.2 Climate

The county has a sub-tropical monsoon climate with a mean temperature of 16.4°C, and an annual mean precipitation of 1,421mm. The temperatures are negatively correlated with altitude in the county. The annual mean temperature is around 11.6 °C - 12.8°C above the altitude of 1,800 metres, and 17°C - 21°C below 800 metres. Figure 4.2 shows the long-term monthly mean temperature in the highlands (the county meteorological observatory was located at the altitude of 1,543 metres above sea level before September 1997) from 1986 to 1995. The monthly mean temperature in the lowlands (the meteorological observatory was moved to the lowland at an altitude of 265 metres at the end of 1997) from 1998 to1999 is also shown in Figure 4.2. The altitude difference of 1,278 of the observatories is reflected in an annual mean temperature difference of 8.3 °C, or a temperature fall of 0.65°C per 100-metre altitude.

There are two seasons in Yuangyang County. The rainy season is from May to October, the dry season from November to April. 80% of the annual rainfall was in the rainy season, and 50% of the annual rainfall occurred from May to July. Nevertheless, the amount of rainfall is not homogeneously distributed in the county. The highlands have much more rain (Figure 4.3) and a much higher humidity than the lowlands (Figure 4.4).

Figure 4.2. The mean temperature in the highlands and low lands in Yuanyang, Yunnan, China

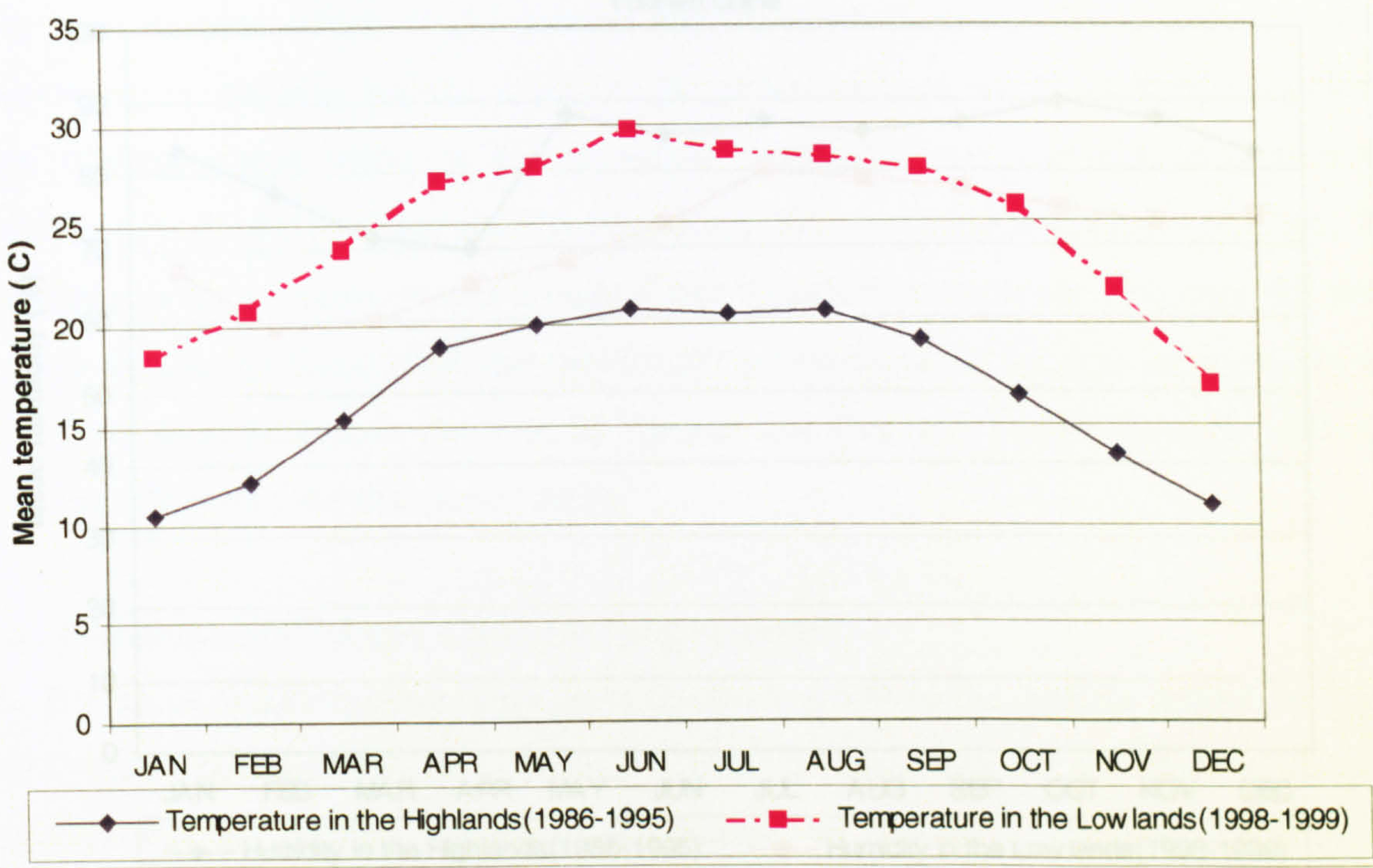


Figure 4.3. The mean rainfall in the highlands and lowlands in Yuangyang, Yunnan, China

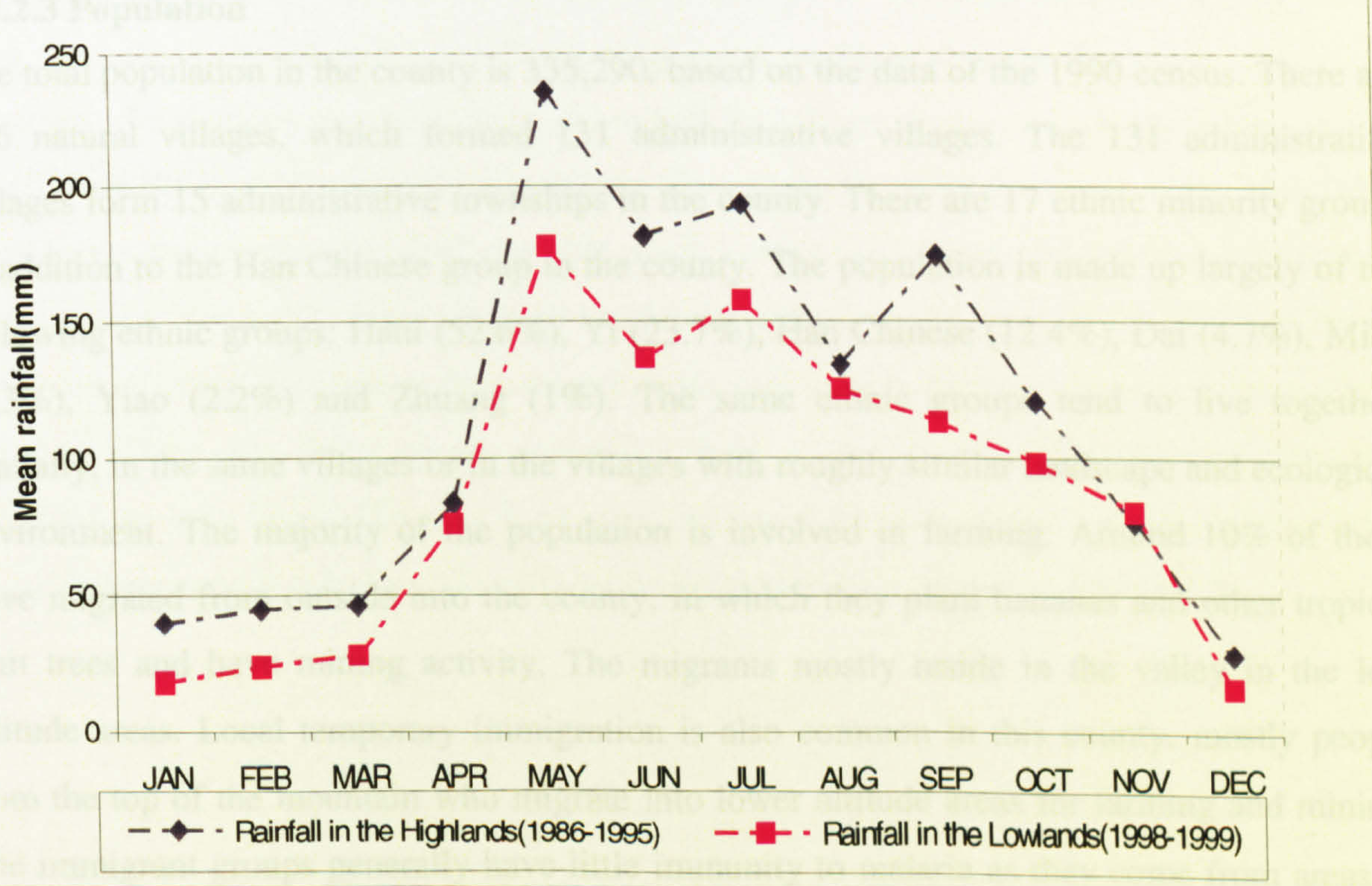
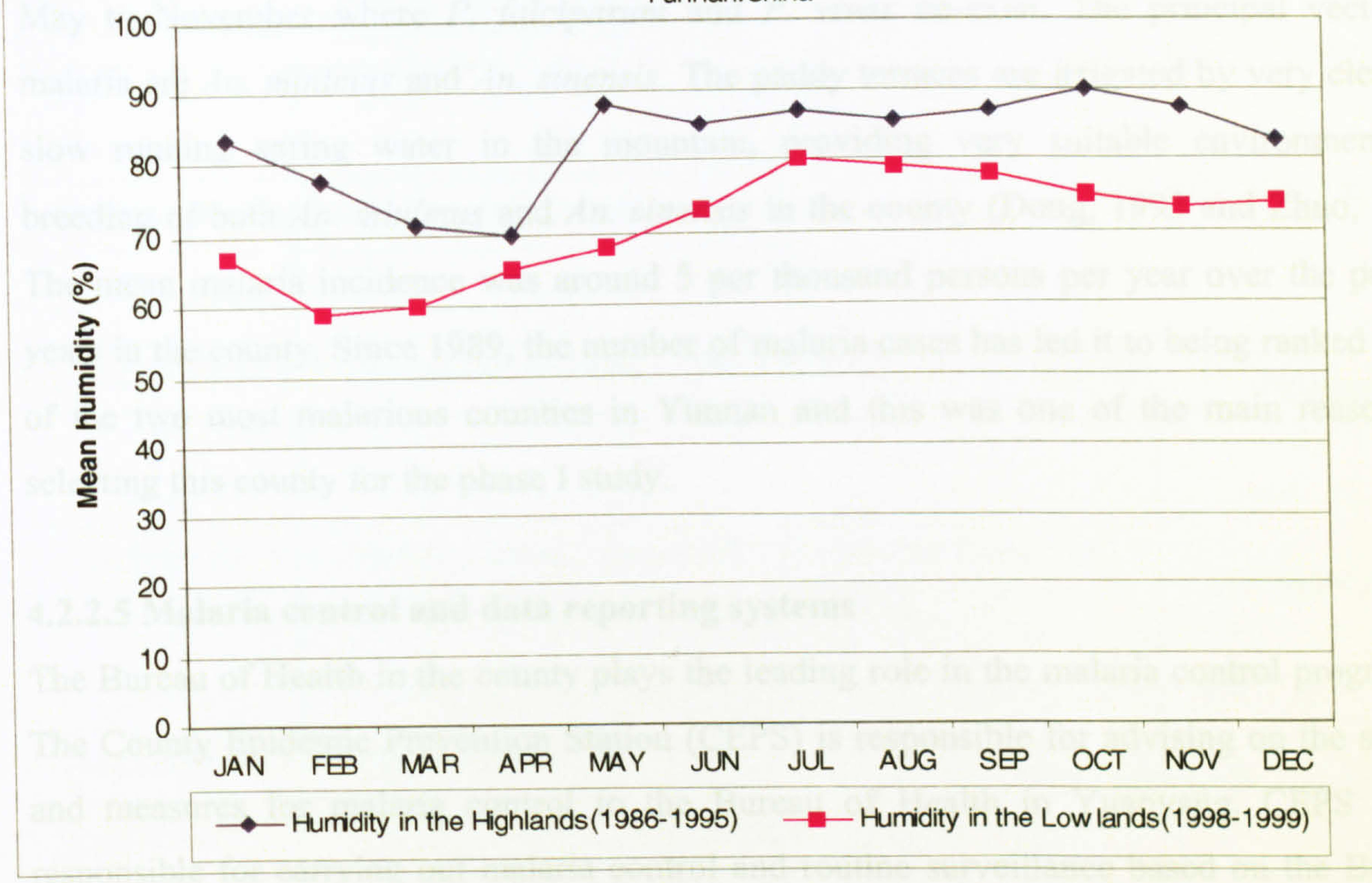


Figure 4.4 The mean relative humidity in the highlands and low lands in Yuanyang, Yunnan.China



4.2.2.3 Population

The total population in the county is 335,290, based on the data of the 1990 census. There are 936 natural villages, which formed 131 administrative villages. The 131 administrative villages form 15 administrative townships in the county. There are 17 ethnic minority groups in addition to the Han Chinese group in the county. The population is made up largely of the following ethnic groups: Hani (52.6%), Yi (23.7%), Han Chinese (12.4%), Dai (4.7%), Miao (3.3%), Yiao (2.2%) and Zhuang (1%). The same ethnic groups tend to live together, spatially, in the same villages or in the villages with roughly similar landscape and ecological environment. The majority of the population is involved in farming. Around 10% of them have migrated from outside into the county, in which they plant bananas and other tropical fruit trees and have mining activity, The migrants mostly reside in the valley in the low altitude areas. Local temporary immigration is also common in this county, mostly people from the top of the mountain who migrate into lower altitude areas for farming and mining. The immigrant groups generally have little immunity to malaria as they come from areas of lower or zero endemicity.

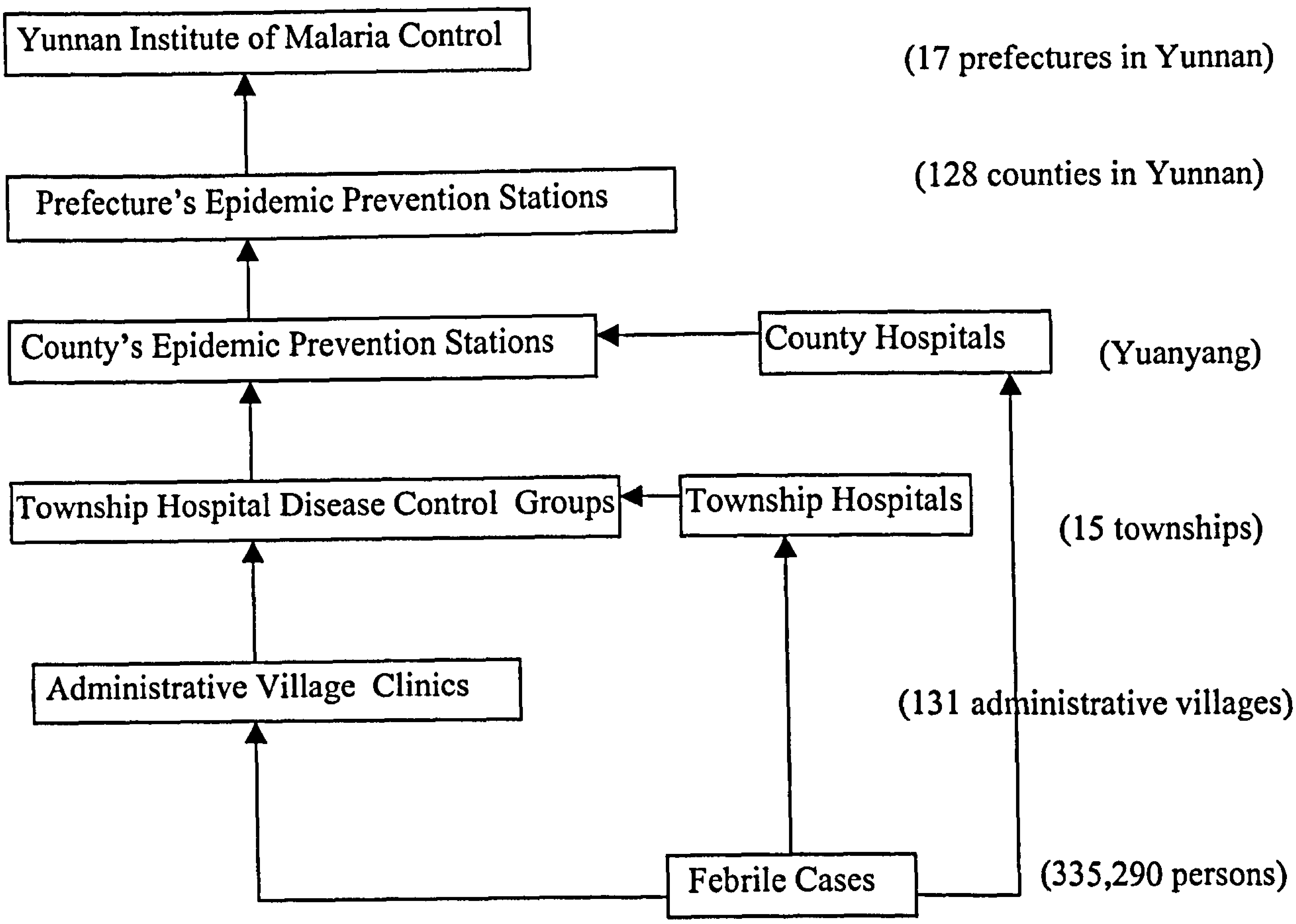
4.2.2.4 Malaria and its vectors

Malaria transmission is perennial in the lower altitude areas, peaking in the rainy season from May to November where *P. falciparum* and *P. vivax* co-exist. The principal vectors of malaria are *An. minimus* and *An. sinensis*. The paddy terraces are irrigated by very clear and slow running spring water in the mountain, providing very suitable environments for breeding of both *An. minimus* and *An. sinensis* in the county (Dong, 1993 and Zhuo, 1991). The mean malaria incidence was around 5 per thousand persons per year over the past ten years in the county. Since 1989, the number of malaria cases has led it to being ranked as one of the two most malarious counties in Yunnan and this was one of the main reasons for selecting this county for the phase I study.

4.2.2.5 Malaria control and data reporting systems

The Bureau of Health in the county plays the leading role in the malaria control programme. The County Epidemic Prevention Station (CEPS) is responsible for advising on the strategy and measures for malaria control to the Bureau of Health in Yuanyang. CEPS is also responsible for carrying out malaria control and routine surveillance based on the Bureau's decision, and for collecting and compiling of routine malaria data. Each township has its own hospital. The disease control and surveillance groups in township hospitals are responsible for the malaria control and surveillance programme in the townships. Their jobs also include reading blood slides that have been submitted by village doctors in village clinics and the staff from the township hospital laboratory. They make the final diagnosis of the blood slides and report malaria data to CEPS monthly. Each village has its own clinic which is responsible for taking a blood slide from whoever has febrile episodes, and submits the blood slides with relevant demographic data to their township hospital weekly in the main malaria transmission season. There are three county level hospitals in the county. The hospital laboratory technicians read blood slides in their own hospital laboratories, and report malaria data to CEPS every month directly. The data reporting systems are shown in Figure 4.5.

Figure 4.5. A paradigm of malaria routine data reporting and surveillance system in Yunnan



4.2.3 Data collection

4.2.3.1 Malaria and population data

The first steps of the data collection involved visiting the CEPS in Yuanyang to collect routine malaria data for the past 10 years (from 1987 to 1996). If the malaria data were not available or not properly recorded in the CEPS, the disease control groups in the relevant administrative township hospitals were visited for further collection, checking and correction. The number of confirmed malaria cases in local permanent residents were collected at the administrative village level from 1987 to 1996. The data on malaria cases from the immigrant population and non-permanent residents were not collected. Although we originally hoped to collect detailed information on all individual cases for demographic and other variables such as, name of patient, age, sex, date of illness , in reality it was not possible to collect so much data within such a short period, and this also applied to later stages of data entry and data analysis in the phase I study. The data on the number of febrile cases examined in the county in previous decades were also collected.

Data for population enumeration in each administrative village, ethnic group distribution, and other demographic data were collected from various government departments in Yuanyang. The population census data of 1990 were also obtained from the county government.

4.2.3.2 Environmental data (maps)

Topographical and terrain maps. The topographical and terrain maps, based on the air photogrammetry survey in 1969 and produced by the Chinese Military Authority in 1978, were collected from Yunnan Provincial Bureau of Land Survey and Cartography. There are six separate map sheets to cover the whole topographic and terrain features of Yuanyang. The scale of the topographical and terrain maps is 1:100,000. The geographical co-ordinate system of the maps is Beijing 1954 Geographical Co-ordinate System with a latitude and longitude grid. The altitude contour system of the topographical maps is based on the Huanhai Altitude System of China, with an elevation contour interval of 40 metres.

Land use map. A land use map (1:150,000) of Yuanyang county was generated by the Bureau of Agriculture of Yuanyang County in 1986. It is based on the results of land use and soil surveys by the County government from 1983 to 1985. The geographical co-ordinate system of the map is the Beijing 1954 Geographical Co-ordinate System with a latitude and longitude grid. The categories were revised based on the characteristics of land use and land cover.

Soil map. A soil map (1:150,000) of Yuanyang County was produced by the Bureau of Agriculture of Yuanyang County in 1986. It was based on the results of the above land use and soil surveys and used the same geographical co-ordinate system as the land use map.

Administrative boundary map. The administrative boundary map (1:100,000) was produced by the Bureau of Transportation of Yuanyang County in 1996. The boundaries were drawn based on the detailed land surveys by the Bureau of Land in the county in 1992. The landscape features and river systems of the map were based on the 1:100,000 topographical and terrain maps produced in 1978. The boundary map comprises county, township and administrative village boundaries. There are no geographical co-ordinate grids on the administrative maps.

Meteorological data (precipitation, temperature, humidity *etc.*) for the past ten years were collected from the meteorological organisation in Honghe Prefecture Meteorological Bureau and Yuanyang Meteorological Observatory. The climate data were on a monthly basis. Meteorological maps were also obtained in Yunnan Provincial Meteorological Bureau, However, the scales of the maps were small as they were provincial maps.

4.3 Data management and analysis

4.3.1 Definition of malaria

The criterion of confirmed malaria diagnosis is "a history of recent fever case with malaria parasites found on microscopic examination of thin and/or thick blood slides". The parasitaemia-positive cases with no fever, detected in an active parasitaemia survey, were not included in the analysis.

4.3.2 Map digitisation and cleaning

All thematic maps were traced from the original map sheets onto digitisation sheets (tracing pad of Mylar) prior to digitisation. They were digitised by using Summa graphics and the ARC/INFO software package of ADS module. All digitisation sheets were registered to a master tic file (with a minimum of eight tic positions for each sheet). The accuracy of the digitisation was assessed using the Root Mean Square error (RMS), which measures the errors between the new digitising session and a previous one (how well they were matched). RMS errors of 0.002 were acceptable when new data was digitised into existing coverage in the present study. If this RMS value was exceeded the tics were re-digitised. Because no geographical co-ordinates were on the administrative boundary sheet, 10 intersection points of streams (identifiable from both the administrative map and the topographical and terrain maps) were used as tic points for the administrative boundary map.

After digitising, all coverages were constructed with their topology in ARC. Errors were then corrected in ARCEDIT and the final coverage then cleaned in ARC. And then all the coverages were geometrically corrected by using known tic co-ordinates (in latitude/longitude). To carry out spatial analysis among the landscape and environmental maps, the spheroid angle co-ordinates (latitude/longitude) have to be projected onto a two-dimensional, flat surface planar co-ordinate system. Therefore, the spheroid co-ordinates (latitude and longitude) of all coverages (maps) in the present data base were then projected onto a planar co-ordinate system, the Universal Transverse Mercator (UTM) co-ordinate system in ARC/INFO. UTM preserved the shape of the area, the distortion of area is relatively small, and it works best for relatively small areas such as the present study area (DeMers, 1997). The attributes of the spatial features were numerically coded and appended to relevant polygons and/or arcs attribute tables manually in ARCEDIT and ARCPLOT.

The terrain maps of Yuanyang County consist of six separate map sheets. A 200-metre interval between contours was digitised. The six-sheets were digitised individually and were merged into one single map sheet to form the topographical and terrain features of Yuangyang county in ARCEDIT. The terrain maps were cleaned as polygons. The areas (polygons) between two contours were coded as being at the mean altitude of the two contours. The attribute of mean altitude of the polygon (terrain map) was appended to the relevant polygon attribute tables manually in ARCEDIT and ARCPLOT. The administrative village boundary map was overlain with a terrain map coded as polygons in ARCPLOT. The mean altitude of each administrative village was calculated according to all the terrain polygons weighted by the respective areas. The administrative village boundary maps were overlain with land-use and soil maps to calculate the proportion of area that each landscape element occupied within each administrative village.

4.3.3 Spatial and statistical analysis

Malaria and population data were entered into an Excel spread-sheet. After re-checking and cleaning, the data were transferred into ASCII file format, and then attached onto the INFO file in ARC/INFO using a common identification number. The data were also transferred into STATA file format for further statistical analysis.

Malaria incidence rates were calculated for each individual administrative village based on routine malaria data. The numerator of the incidence rate was the number of the confirmed malaria cases (episodes) reported during the year. Mixed infections were counted both as *P. vivax* and *P. falciparum* infections. All malaria episodes were counted separately, whether they occurred in the same person or not, provided they were distinct from each other (episodes at time intervals of 28 days or longer were considered to be separate cases, if they had taken full dosages of anti-malarials). The denominator of the incidence rate is the population in the administrative village for the year. Malaria incidence rates were calculated for each individual administrative village during the 10-year period. As the malaria incidence rates were not normally distributed, they were transformed onto a logarithmic scale after adding one [Log (x + 1)] and geometric mean incidence rates were calculated in each administrative village for the past 10 years by taking the antilog of the mean logged data [ΣLog(x + 1)/n], and then subtracting 1. The geometric mean incidence rates of *P. falciparum,* and *P. vivax* were input into ARC/INFO in each individual administrative village coded area. Two malaria coverages, for *P. falciparum* and for *P. vivax*, were generated in the ARC/INFO environment accordingly. Malaria and population are assumed to be distributed homogeneously within an administrative village.

The data analysis used the display and integration capability of GIS, showing malaria distribution against the coverages of various landscape and environmental variables. The statistical analysis was based on the administrative village level. The landscape and environmental variables of each administrative village were obtained through overlaying the boundary map with the topographical maps, land use map and soil map. Geometric malaria incidence rates in each administrative village were used as dependent variables, the landscape and environmental variables were used as independent variables. Prior to statistical analysis, the landscape proportions of administrative villages were subjected to an angular transformation, as is required to normalise proportional and percentage data prior to analysis (Armitage & Berry, 1987). These transformed elements were subsequently used as the variables in statistical analysis. Multiple logistic regression was used to discriminate the importance of different variables in relation to the high and low incidences of *P. vivax* in administrative villages, and the present or absence of *P. falciparum* in the administrative villages. Crude odds ratios were estimated from univariate models and adjusted odd ratios were estimated after controlling for confounding effects of other variables. The analyses were carried out by using the software package STATA. To cope with spatial autocorrelation at administrative township level, multilevel logistic regression was used for the further analysis. The detailed methodology will be described in section 4.4.4 of this Chapter.

4.4 Results

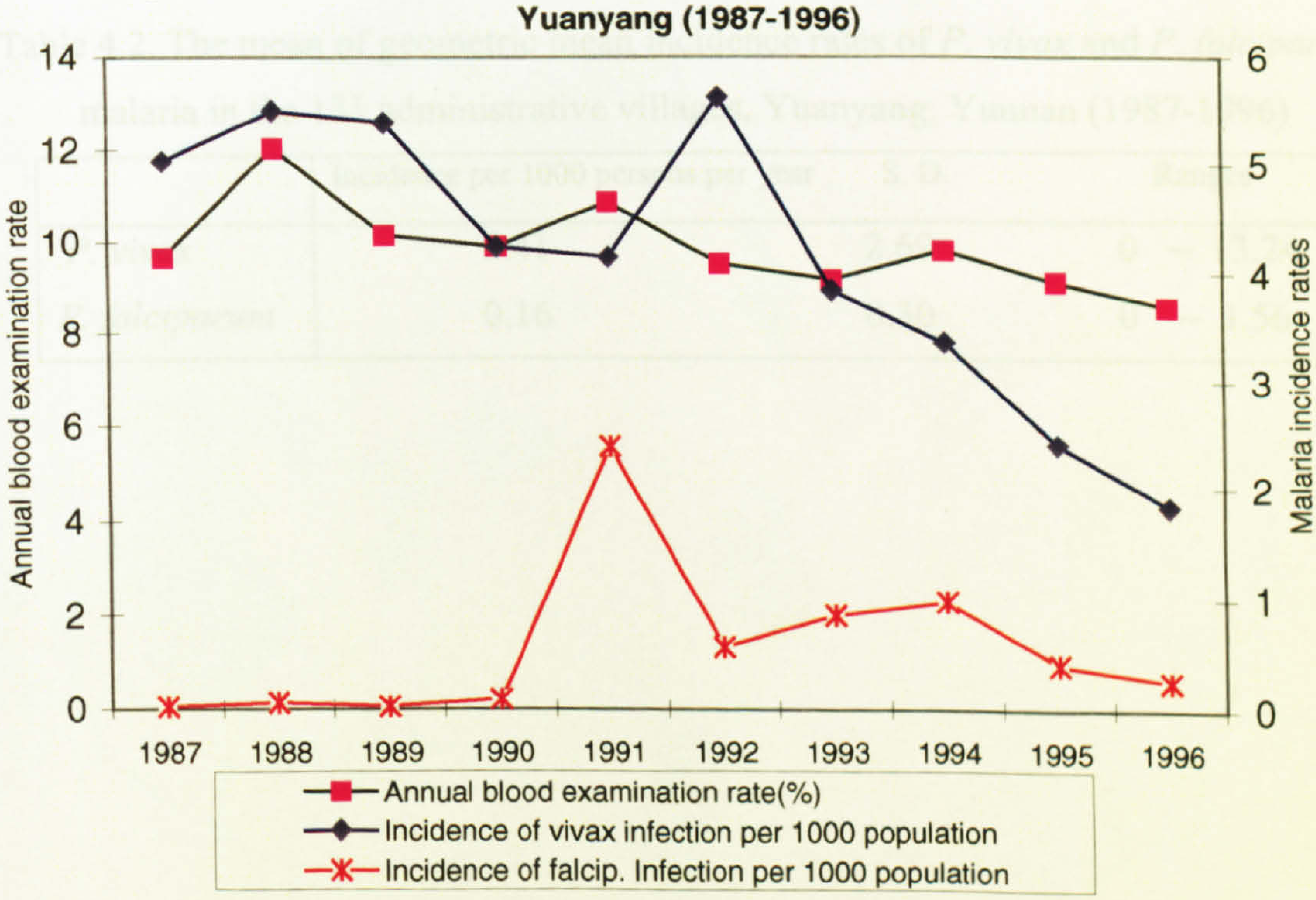
4.4.1 Malaria and environmental variables

The population in Yuanyang was around 330,000 as shown in Table 4.1. During the 10-year period (1987-1996), a total of 330,796 blood slides from febrile patients was collected and examined among local permanent residents in Yuanyang County. The proportion of annual blood slide examinations for febrile patients ranged from 8.62% in 1996 to 12.04% of the population in 1988. The mean percentage annually of blood examinations was 9.88 over the 10-year period. The detailed results are shown in Table 4.1. From 1987 to 1996, a total of 13,768 *P. vivax* and 1,874 *P. falciparum* and 11 mixed *P.vivax* and *P.falciparum* infection confirmed malaria cases were reported among local permanent residents. The overall mean incidence rates for *P. vivax* and *P. falciparum* infection were 4.2 and 0.6 per thousand person per year, respectively (mixed infection counted as both as *P. vivax* and *P. falciparum* infection) during the 10-year period. The annual variations of malaria incidence rates were obvious during the 10-year period (Figure 4.6), particularly for *P. falciparum*. The number of malaria cases and malaria incidence rates peaked in 1991 and 1992, when there was a *P. falciparum* malaria outbreak in Nujiezai Township in the county (Yang *et al.*, 1992). There has been a trend of decreasing malaria incidence for both *P. vivax* and *P. falciparum* infection in the recent years as indicated in Figure 4.6, which might be due to strengthening of the malaria control measures after the malaria outbreak in Nujiezai (He Yuanchin, personal communication). To ensure a relatively homogenous socio-economic background of the study population and to reduce migration bias, the populations in small towns and the capital city of Yuanyang were excluded from both the numerator and denominator for further analysis. Since the people in the small towns and the capital city have a better socio-economic situation and are more mobile for business, it would be very difficult to identify where they obtained their infection. Consequently, 2,537 *P. vivax* and 493 *P. falciparum* malaria cases were excluded from present analysis because those cases came from the urban areas.

Table 4.1. Malaria and annual blood examination rate in Yuanyang (1987-1996)

Year	No. pop.	No. slide	%	No. *P.vivax*	No. *P. falciparum*	No. mixed
1987	317,822	30,694	9.66	1,602	9	0
1988	321,136	38,674	12.04	1,766	21	0
1989	325,690	33,100	9.84	1,755	10	0
1990	335,290	33,243	9.91	1,424	32	0
1991	336,223	36,547	10.87	1,391	797	4
1992	339,893	32,406	9.53	1,911	193	0
1993	342,171	31,474	9.20	1,316	249	3
1994	343,817	33,794	9.83	1,152	335	4
1995	339,113	31,081	9.17	817	139	0
1996	345,357	29,783	8.62	634	89	0
Total	3,346,572	330,796	9.88	13,768	1,874	11

Figure 4.6. Annual blood examination rates and malaria incidence rates in Yuanyang (1987-1996)



Of the 131 administrative villages, 8 (6%) had no *P. vivax* case reported during the 10-year period; 70 (53%) had no *P. faciparum* cases; and 78 (60%) had geometric mean incidence rates for *P. vivax* infection equal to or lower than 2 per thousand population during the 10 year period. The 10-year geometric mean incidence rates of *P. vivax* infection ranged from 0 (8 administrative villages) to 13.24 per thousand persons, and of *P. falciparum* from 0 (70) to 1.56 per thousand persons. The overall geometric mean incidence rates for *P. vivax*, and *P. falciparum* of the 131 administrative village were 2.41, and 0.16 per thousand persons per year in the county during the 10-year period, respectively (Table 4.2). The spatial patterns of *P. vivax* and *P. falciparum* malaria incidence rates over the 10-year period are shown in Figures 4.7 and 4.8, respectively. As the Figures 4.7 and 4.8 indicate, malaria for both *P. vivax* and *P. falciparum* was highly variable within Yuanyang County. It was mainly distributed along the Red River. The ranked distribution of the geometric mean incidence rates of the 131 administrative villages in the County from 1987 to 1996 for *P. falciparum* and *P. vivax* are shown in Figures 4.9 and 4.10, respectively.

Table 4.2. The mean of geometric mean incidence rates of *P. vivax* and *P. falciparum* malaria in the 131 administrative villages, Yuanyang, Yunnan (1987-1996)

	Incidence per 1000 persons per year	S. D.	Ranges
P. vivax	2.41	2.69	0 ~ 13.24
P. falciparum	0.16	0.30	0 ~ 1.56

Figure 4.7. The spatial distribution of P. vivax in Yuanyang, Yunnan (1987-1996)

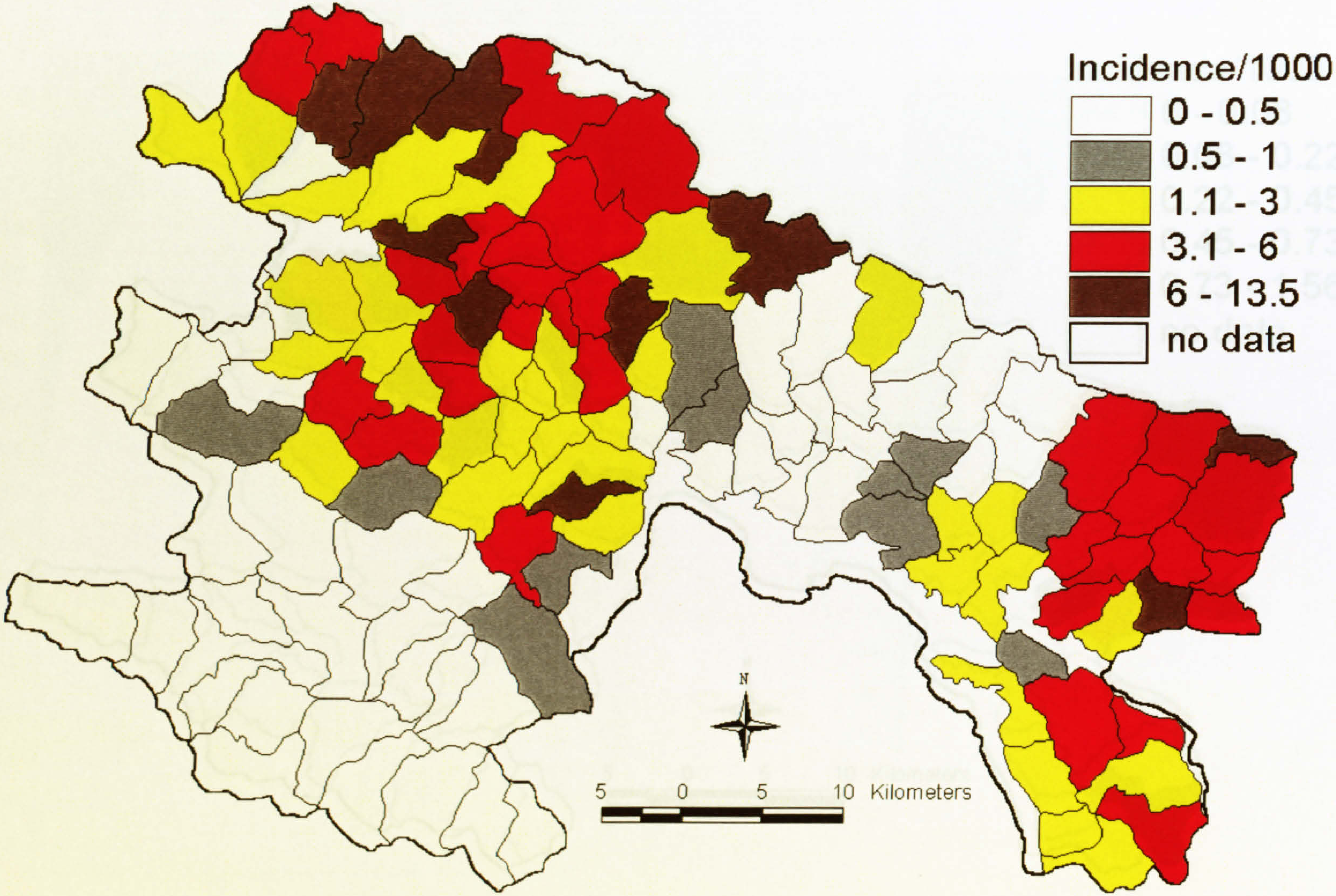


Figure 4.8. The spatial distribution of P. falciparum in Yuanyang, Yunnan (1987-1996)

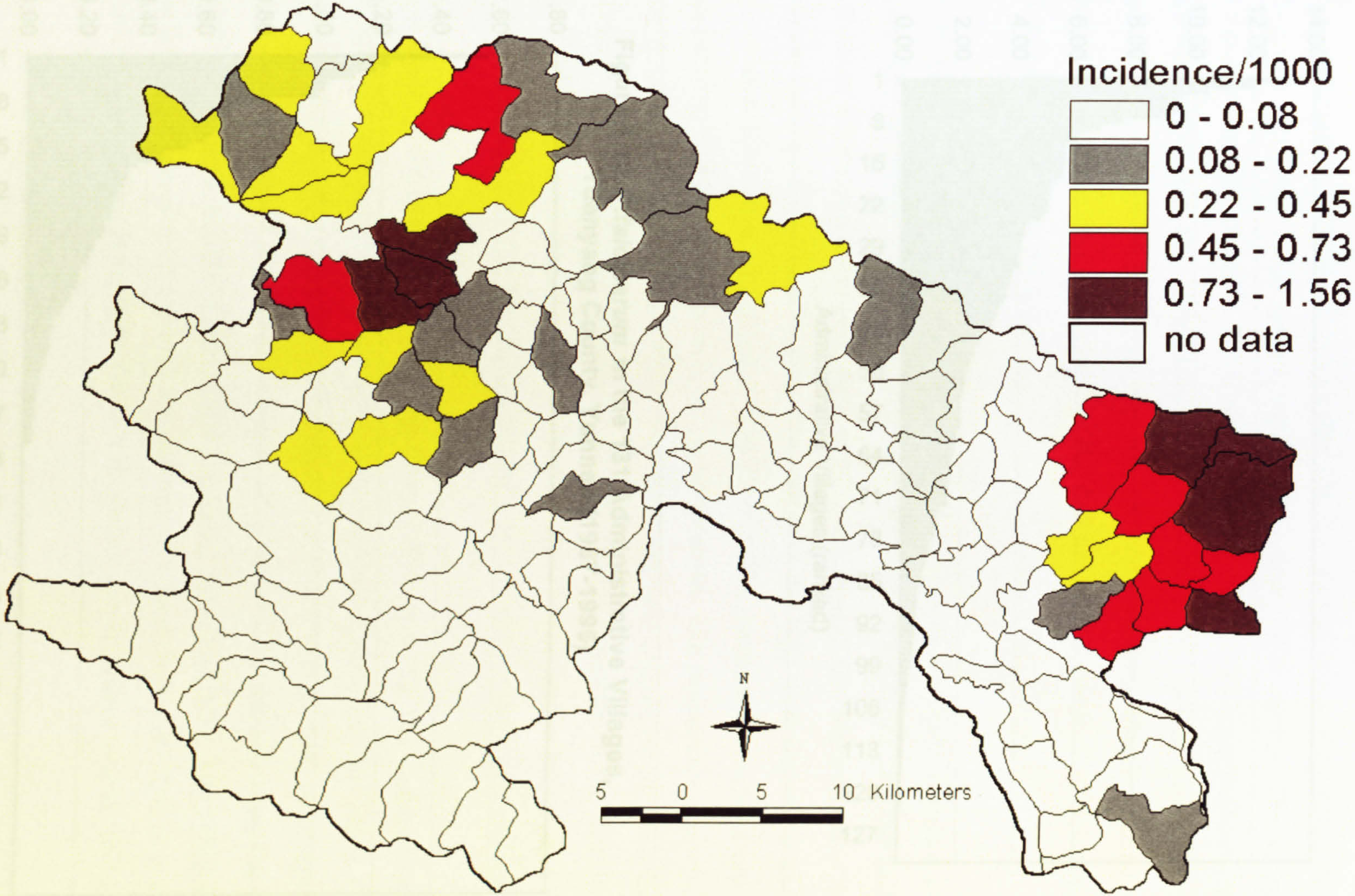


Figure 4.9. *P.vivax* in the 131 Administrative Villages, Yuanyang County, Yunnan (1987-1996)

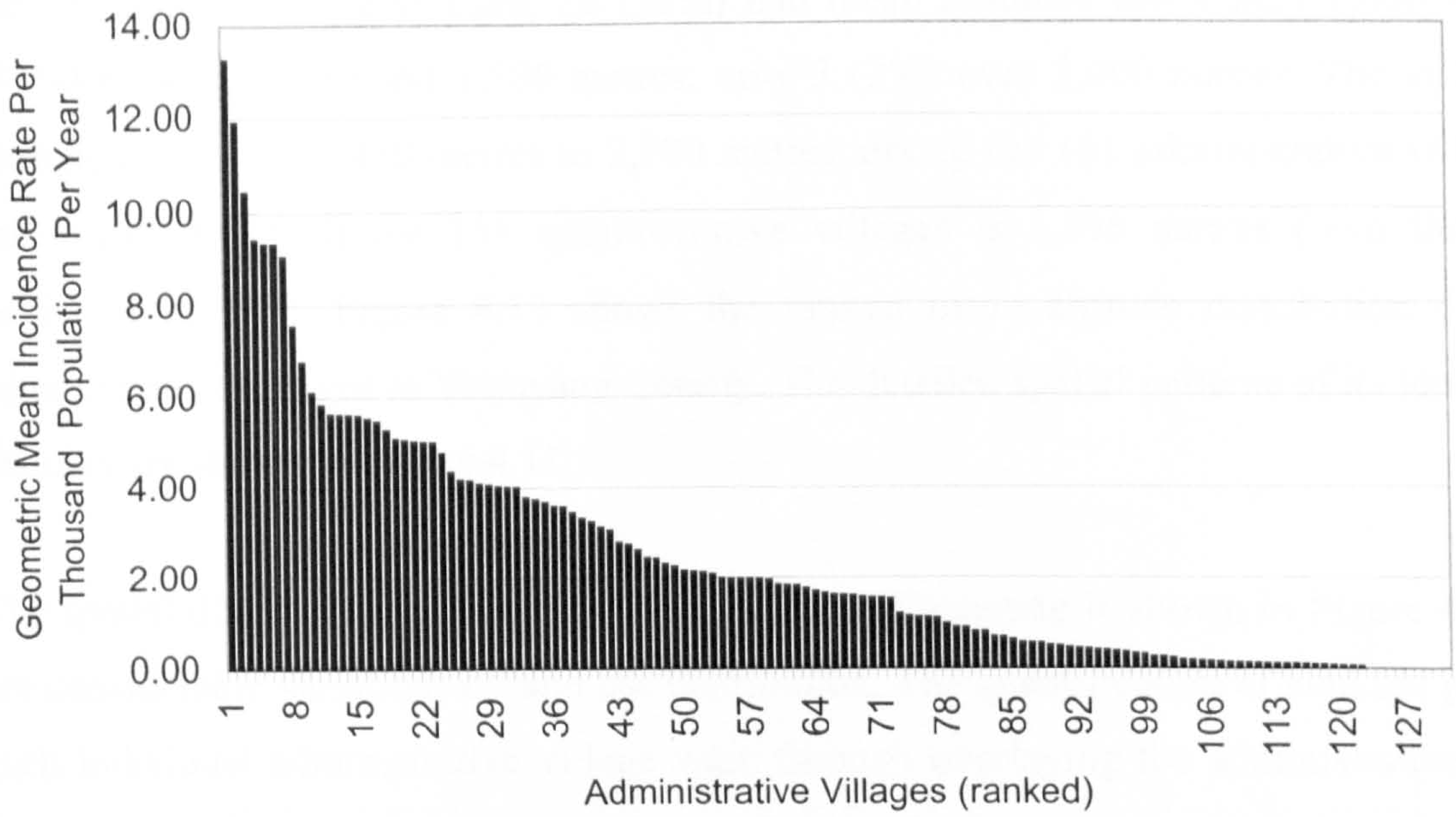
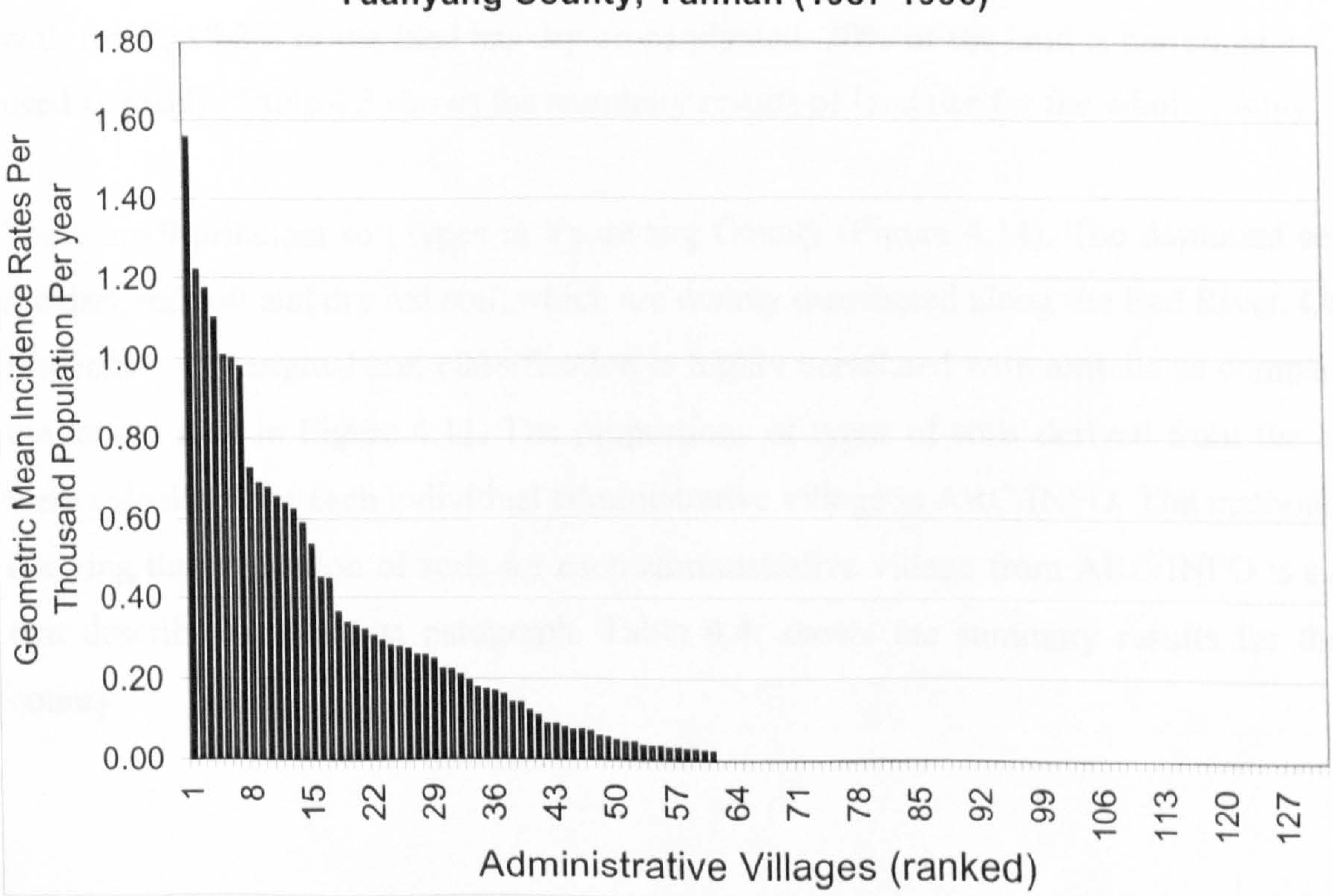


Figure 4.10. *P.falciparum* in the 131 Administrative Villages, Yuanyang County, Yunnan (1987-1996)



There are remarkable variations of the terrain in Yuanyang County (Figure 4.11). The lowest place is at an altitude of 144 metres and the highest 2,939 metres above the sea level. The mean altitude for each administrative village in the county was derived using ARC/INFO. Of the 131 administrative villages, 20 (15%) had mean altitudes lower than 1,000 metres; 58 (44%) between 1,000 and 1,500 metres; only 3 (2%) over 2,000 metres. The village mean altitude ranged from 480 metres to 2,100 metres among the 131 administrative villages. The mean altitude of all the 131 administrative villages is 1,396 metres (with the standard deviation of 361). Figure 4.12 shows the ranked mean altitude distribution of the 131 administrative villages in Yuanyang County. The detailed spatial patterns of its topographical features are shown in Figure 4.11.

The spatial distribution of land use patterns in the Yuanyang is shown in Figure 4.13. There are considerable variations of land use distribution. The quantification of land use patterns for each individual administrative village were through overlaying the administrative boundary map with the land use map. The total areas for each land use type were derived from the land use map for each administrative village in ARC/INFO. The database of the land use types of all administrative villages was exported into a spreadsheet. The proportion of each land use type over the total area within the administrative village was calculated for each individual administrative village for further statistical analysis. Around a quarter of the land is covered with forest, 10.2% of the land has dry crops planted, 20% of the land is barren, and 13.5% is used as paddy. Table 4.3 shows the summary results of land use for the whole county.

There are 9 principal soil types in Yuanyang County (Figure 4.14). The dominant soil types are dark red soil and dry red soil, which are mainly distributed along the Red River. On visual inspection, the original soil classification is highly correlated with altitude as compared with the terrain map in Figure 4.11. The proportions of types of soils derived from the soil map were calculated for each individual administrative village in ARC/INFO. The methodology of deriving the proportion of soils for each administrative village from ARC/INFO is similar to that described in the last paragraph. Table 4.4. shows the summary results for the whole county

Figure 4.11. Topographical and terrain features in Yuanyang, Yunnan

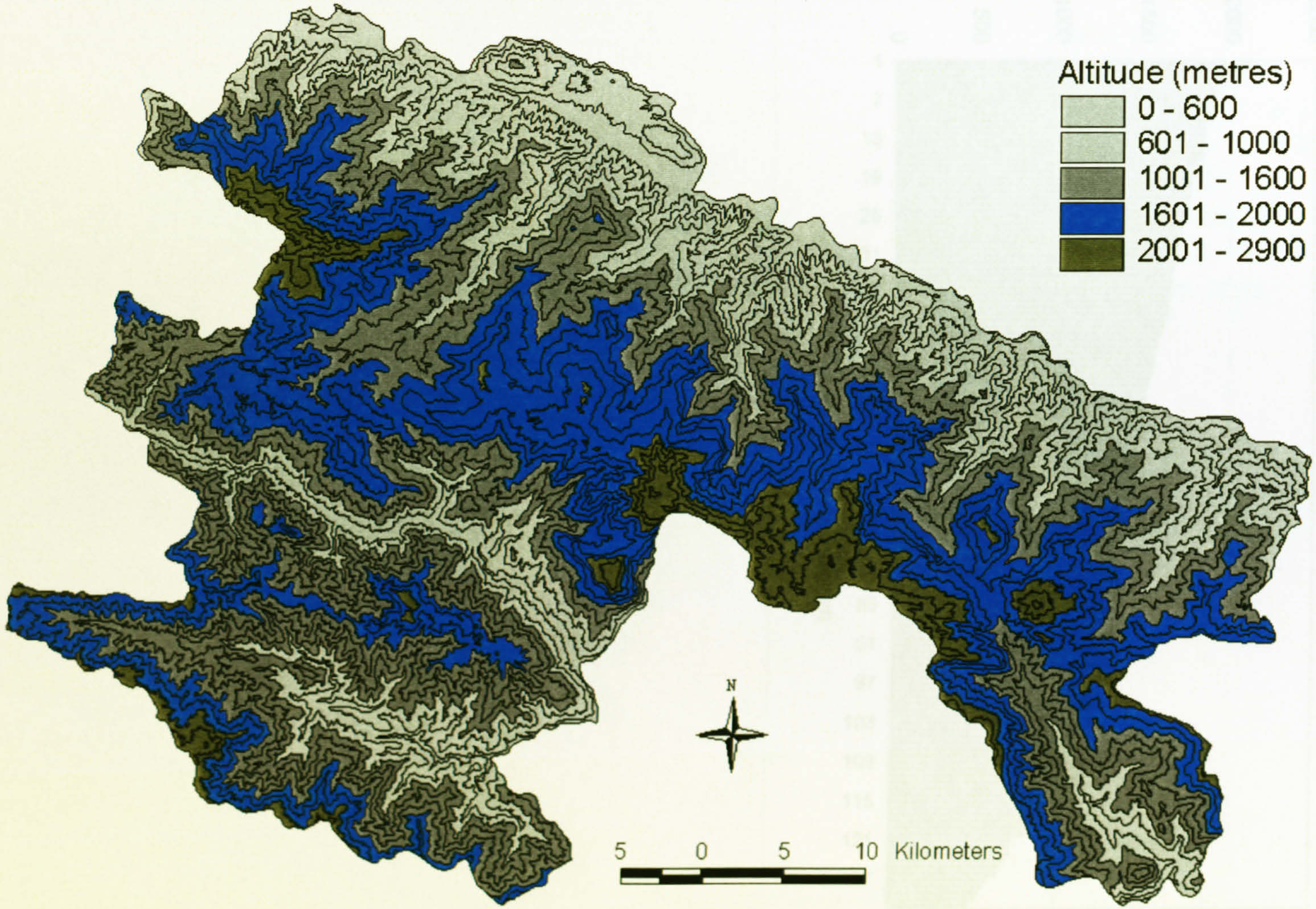


Figure 4.12. The mean altitudes of the 131 Administrative Villages, Yuanyang County, Yunnan

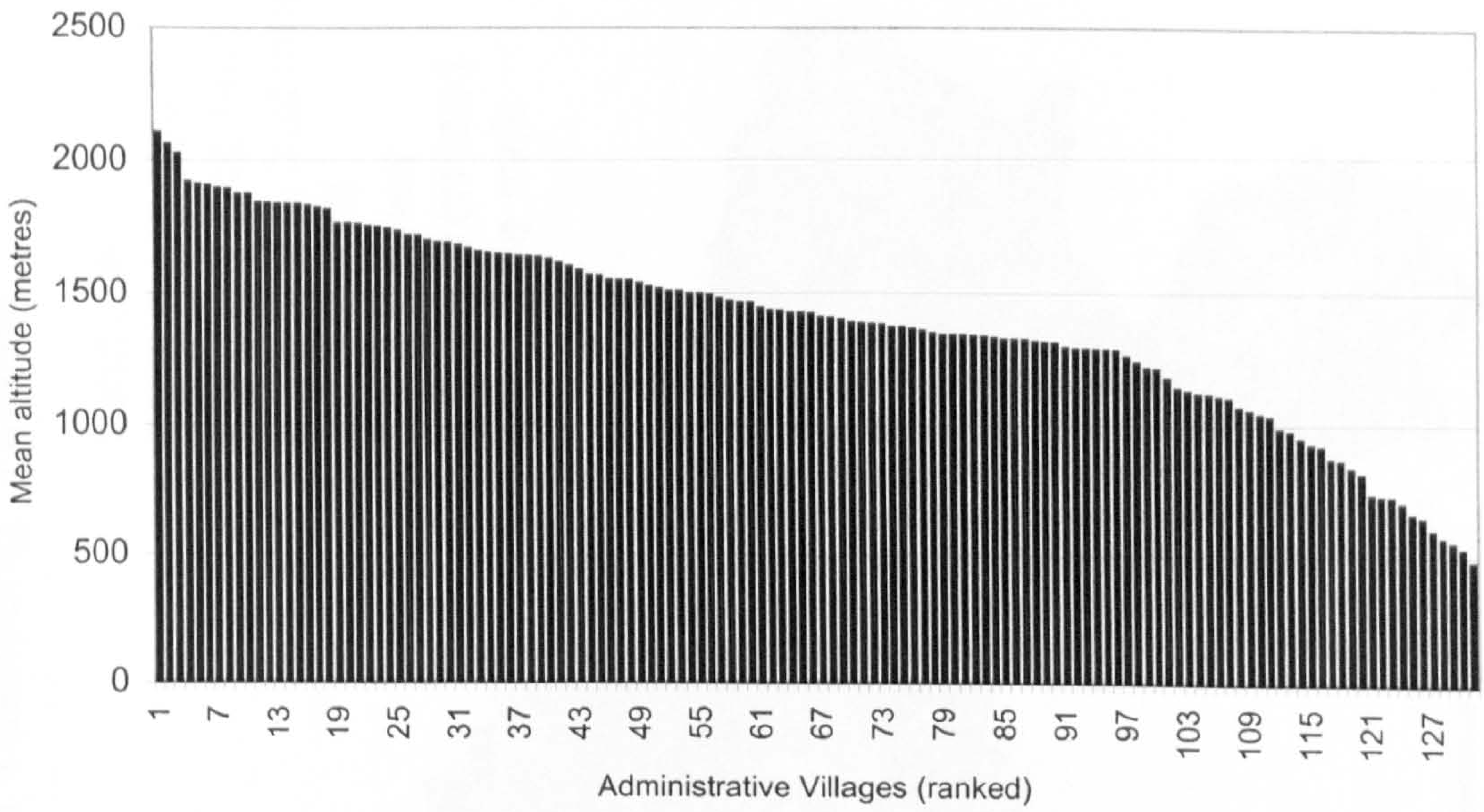


Figure 4.13. The land use and land cover in Yuanyang, Yunnan

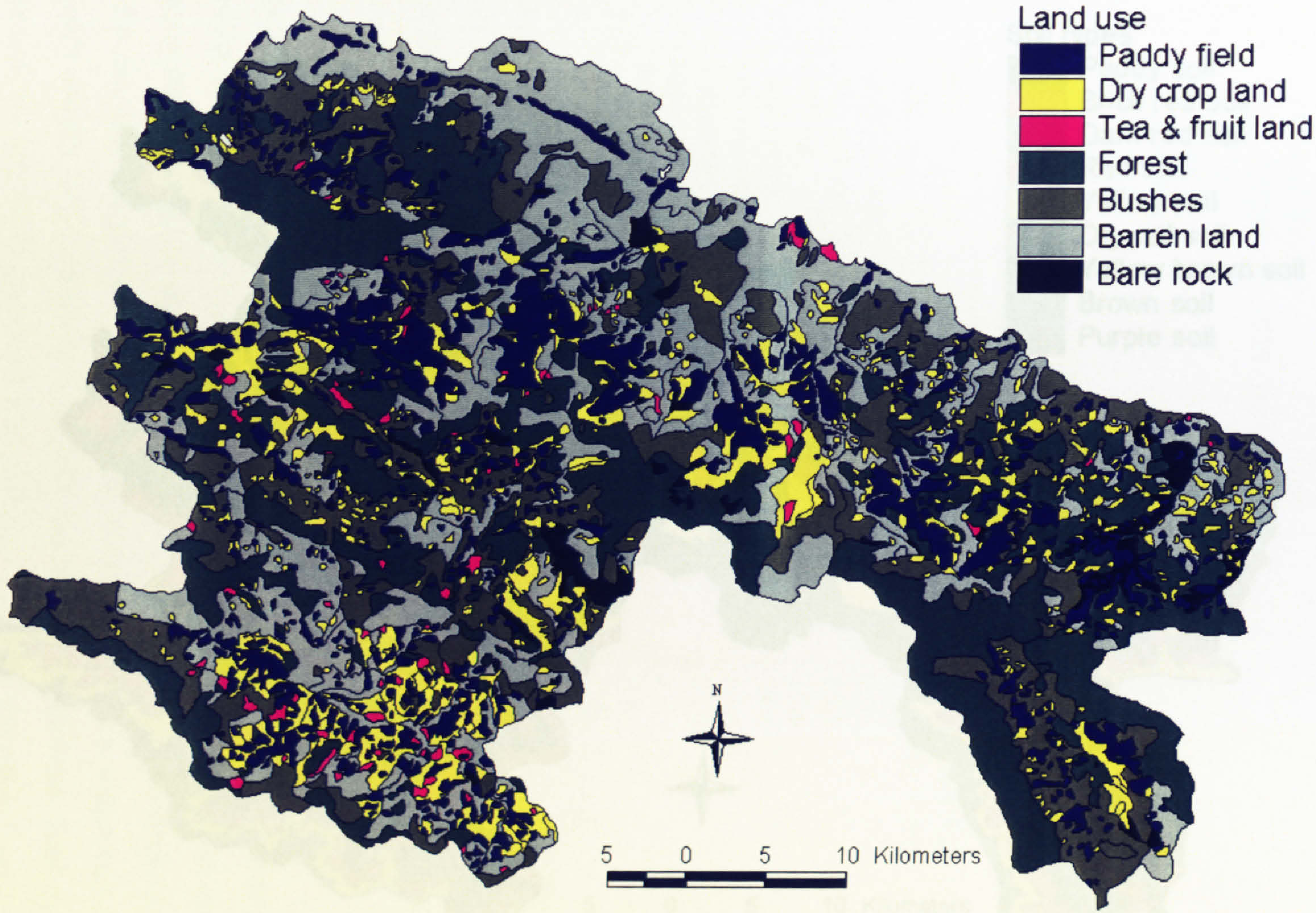


Figure 4.14. The distribution of soils in Yuanyang, Yunnan

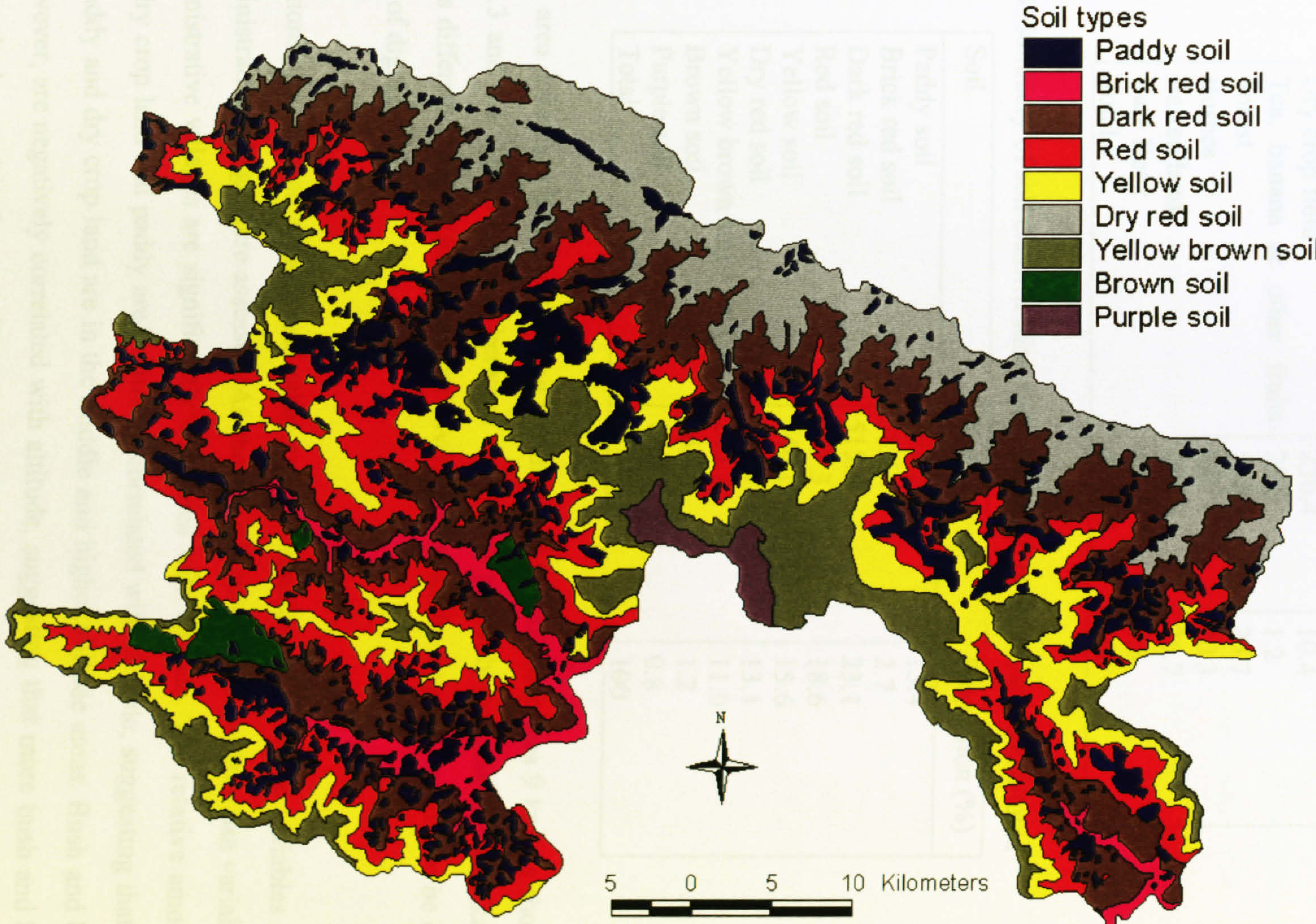


Table 4.3. Summary of land use classification and proportion of Yuanyang covered

Land use element	Km^2	Proportion (%)
Paddy field	295.9	13.5
Dry crop field	230.4	10.4
Tea, banana & other fruits	25.5	1.2
Forest	563.5	25.7
Bushes	446.0	20.3
Barren land	585.3	26.7
Bare Rock	17.9	0.8
Total	2193.9	100

Table 4.4. Summary of soil classification and proportion of Yuanyang constituted

Soil	Km^2	Proportion (%)
Paddy soil	292.2	13.4
Brick red soil	58.6	2.7
Dark red soil	515.2	23.1
Red soil	407.3	18.6
Yellow soil	340.1	15.6
Dry red soil	285.3	13.1
Yellow brown soil	240.7	11.0
Brown soil	26.9	1.2
Purple soil	17.9	0.8
Total	2184.2	100

The total area difference between the land use map and the soil map was 9 km^2 as shown in Tables 4.3 and 4.4. The difference between the two is probably because the shrinkage of maps was different in the different storage environments of the maps. It could also be due to the error of digitisation.

The relationships between altitude and other angular transformed land use variables in the 131 administrative village were assessed. As shown in Figure 4.15 most land use variables in the administrative villages are significantly correlated with altitude. The relative amount of forests, dry crop land and paddy are positively correlated with altitude, suggesting that more forest, paddy and dry crop land are in the middle and higher altitude areas. Bush and barren land, however, are negatively correlated with altitude, suggesting that more bush and barren land are in the lower altitude areas (Table 4.5). The relative amounts of paddy and forest are also found to be correlated with barren land. The correlation matrix for altitude and soil variables in the 131 administrative villages is shown in Table 4.6. The distributions of paddy soil, red soil, yellow soil and yellow brown soil are significantly positively correlated with altitude, suggesting that four soils are mainly distributed in the middle and higher altitude area. On the other hand, the distribution of brick red soil, dark red soil and dry red soil are significantly negatively correlated with altitude, suggesting that these three soils are mainly distributed in the relatively lower altitude area. The correlation matrix confirms the visual appraisal that originally the soil classification was highly dependent on an altitude model. Besides, inter-correlation among soils was also observed (Table 4.6). The strong autocorrelation among landscape and land use variables might make it very difficult to assess the effect of the variables on the risk of malaria in an univariate analysis.

Figure 4.15 The association of altitude with land use variables in 131 administrative village in Yuanyang

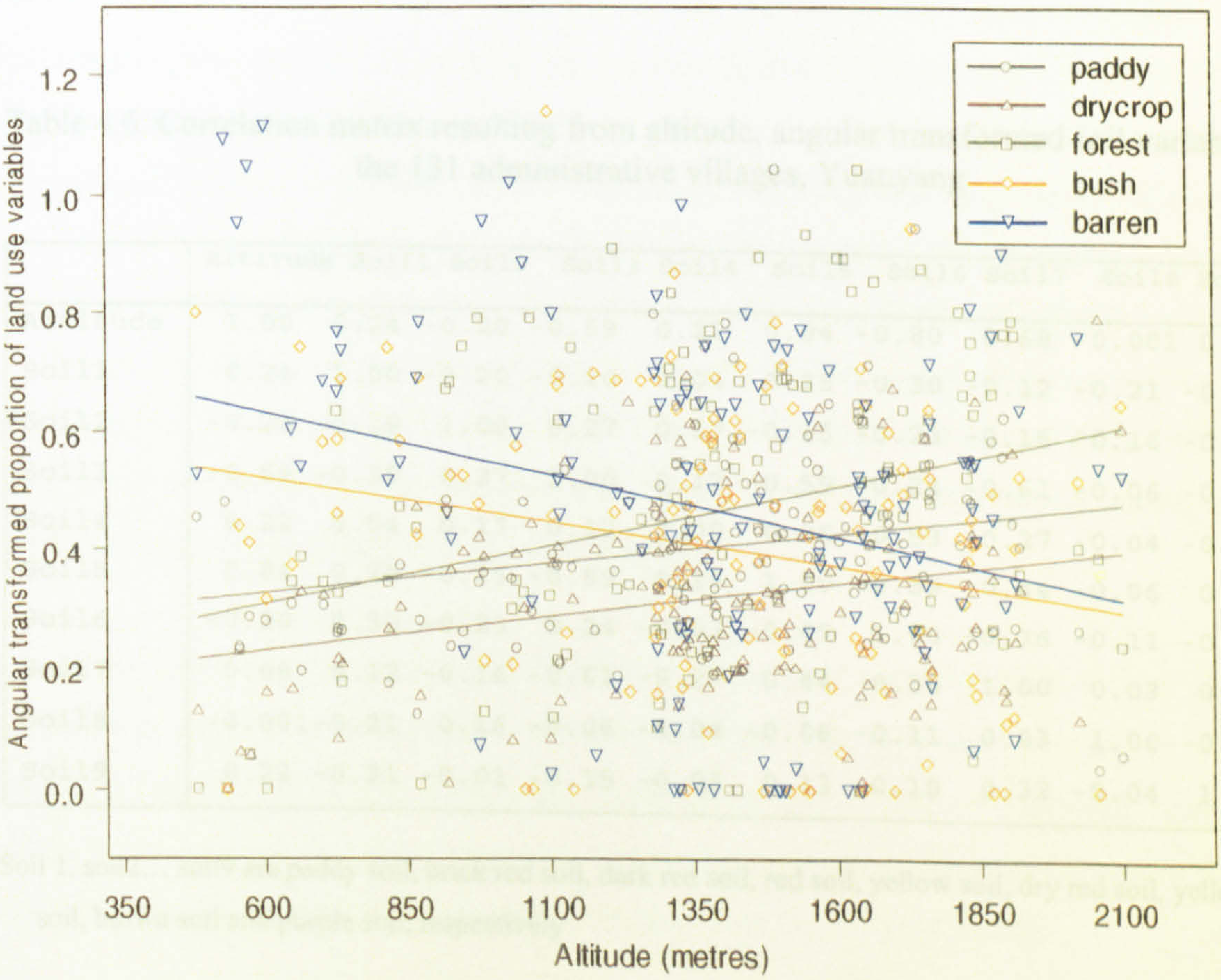


Table 4.5. Correlation matrix resulting from altitude, angular transformed land use variables in the 131 administrative villages, Yuanyang

	Altitude	Paddy	Drycrop	Tea	Forest	Bush	Barren	Rock
Altitude	1.00	0.23	0.26	0.09	0.29	-0.22	-0.28	-0.08
Paddy	0.23	1.00	-0.01	0.03	-0.13	-0.33	-0.18	-0.23
Drycrop	0.26	-0.01	1.00	0.27	-0.18	-0.23	-0.07	0.03
Tea	0.09	0.03	0.27	1.00	-0.16	-0.20	0.10	-0.10
Forest	0.29	-0.13	-0.18	-0.16	1.00	-0.15	-0.61	0.10
Bush	-0.22	-0.33	-0.23	-0.20	-0.15	1.00	-0.26	-0.06
Barren	-0.28	-0.18	-0.07	0.10	-0.61	-0.26	1.00	-0.12
Rock	-0.08	-0.23	0.03	-0.10	0.10	-0.06	-0.12	1.00

Table 4.6. Correlation matrix resulting from altitude, angular transformed soil variables* in the 131 administrative villages, Yuanyang

	Altitude	Soil1	Soil2	Soil3	Soil4	Soil5	Soil6	Soil7	Soil8	Soil9
Altitude	1.00	0.24	-0.20	-0.59	0.22	0.84	-0.80	0.68	-0.001	0.22
Soil1	0.24	1.00	-0.20	-0.20	0.04	0.25	-0.30	-0.12	-0.21	-0.21
Soil2	-0.20	-0.19	1.00	0.27	0.17	-0.15	-0.23	-0.16	0.16	-0.01
Soil3	-0.59	-0.20	0.27	1.00	0.12	-0.59	0.24	-0.61	-0.06	-0.15
Soil4	0.22	0.04	0.17	0.12	1.00	0.26	-0.53	-0.27	-0.04	-0.03
Soil5	0.84	0.25	-0.15	-0.59	0.26	1.00	-0.69	0.44	-0.06	0.13
Soil6	-0.80	-0.30	-0.23	0.24	-0.53	-0.69	1.00	-0.36	-0.11	-0.10
Soil7	0.68	0.12	-0.16	-0.61	-0.27	0.44	-0.36	1.00	0.03	0.32
Soil8	-0.001	-0.21	0.16	-0.06	-0.04	-0.06	-0.11	0.03	1.00	-0.04
Soil9	0.22	-0.21	-0.01	-0.15	-0.03	0.13	-0.10	0.32	-0.04	1.00

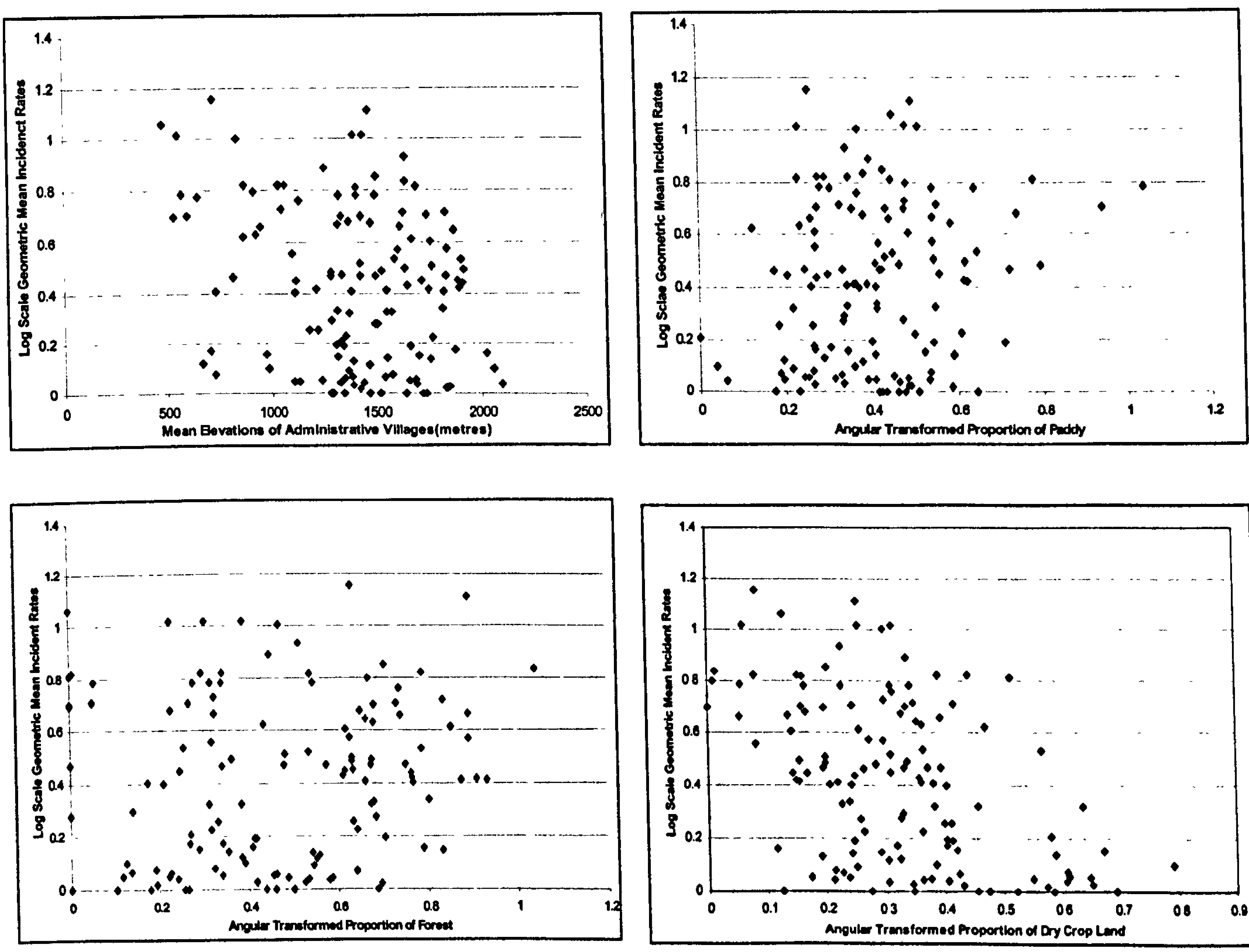
*Soil 1, soil2... soil9 are paddy soil, brick red soil, dark red soil, red soil, yellow soil, dry red soil, yellow brown soil, brown soil and purple soil, respectively

4.4.2 Univariate analysis

The relationship between the risk of *P. vivax* malaria and landscape and environmental variables in Yuanyang was shown by plotting on a scatter diagram the log scale 10-year geometric mean incidence rates against selected landscape and environmental variables (Figure 4.16). The relationships between environmental variables and the risk of *P.vivax* malaria were hardly linear among the majority of the variables, as shown in Figure 4.16. The interpretation of the non-linearity is very difficult. It might be that the relationship between the risk of malaria and potential environmental determinants in the administrative villages is not linear in the study area. Partial autocorrelation among environmental variables makes it much more difficult to use simple linear and non-linear regression to assess the association between the risk of *P. vivax* malaria and the landscape and environmental variables. Furthermore, the variation of malaria reporting in different villages and townships might also confuse the relationship between the risk of malaria and their landscape and environmental indicators. Therefore, the errors of residuals due to reporting and other confounding variables is too big if we use simple linear regression to model the data. Hence, multivariate analyses should be used for further analysis. Nevertheless, the visual appraisals of the log scales risk of *P. vivax* malaria and the environmental variables show a negative correlation between altitude and risk of malaria. Also the administrative villages with more paddy fields tend to have a higher risk of *P.vivax* malaria, and there is a negative correlation between relative amount of dry crop land and risk of vivax malaria.

The associations between the risk of *P. falciparum* malaria and landscape and environmental variables in the 131 administrative villages of Yuanyang were also plotted on the scatter diagrams (Figure 4.17). No significant associations were detected through simple linear statistical analysis or by visual appraisal, due to the majority of the administrative villages in the county having had no *P. falciparum* infection over the 10-year period.

Figure 4.16. The scatter diagram of log scale 10-year geometric mean incidence rates of *P. vivax* malaria[1] and selected landscape and environmental variables[2,3] in the 131 administrative villages, Yuanyang (1987-1996)

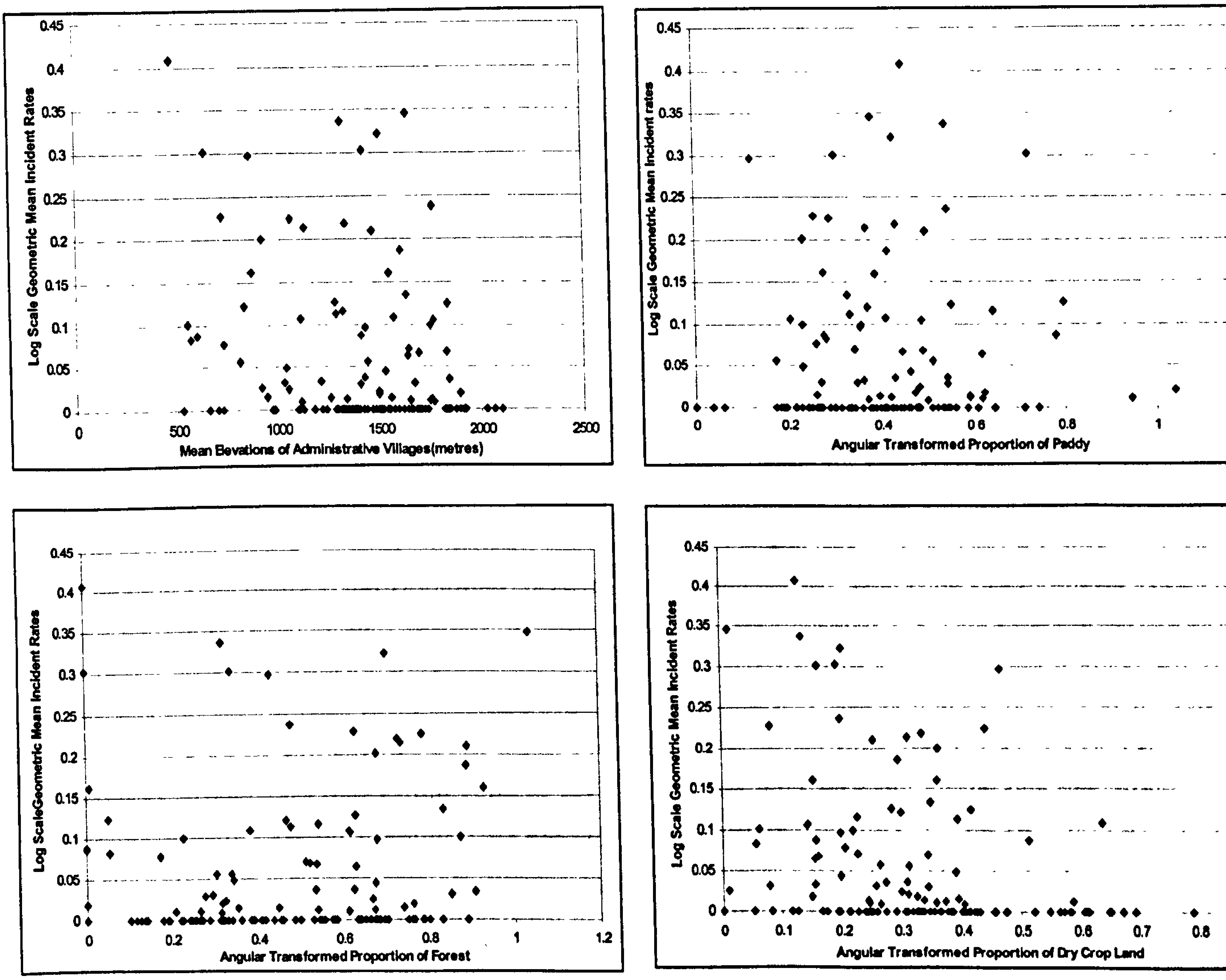


[1] Log geometric mean incidence rate per 1 000 persons per year in the 131administrative villages in Yuanyang

[2] Mean altitude of administrative village (metres)

[3]Angular transformed proportions of the types of land use accounted for in each administrative village

Figure 4.17. The scatter diagram of log scale 10-year geometric mean incidence rates of *P. falciparum* malaria[1] and selected landscape and environmental variables[2,3] in the 131 administrative villages, Yuanyang (1987-1996)



[1] Log geometric mean incidence rate per 1 000 persons per year in the 131 administrative villages in Yuanyang

[2] Mean altitude of administrative village (metres)

[3] Angular transformed proportions of the types of land use accounted for in each administrative village

4.4.3 Results of logistic regression modelling

Multivariate regression is commonly used in the analysis of the determinants of disease distribution, especially for multi-risk phenomena such as malaria. Multivariate linear regression might be not suitable for the present analysis because the relationship between malaria and environmental variables is not linear as shown in Figures 4.16 and 4.17. Poisson regression would be another choice, but the nature of malaria epidemiology in the study area is that the disease is unstable and prone to outbreaks and epidemics. Therefore the data would be unstable, the regression model would give more weight to the area which had an epidemic or outbreak during the data collection period. Hence, logistic regression was used to quantify the risk factors of the administrative villages with higher *P. vivax* incidence versus the villages with the relatively lower incidence, and the villages with *P. falciparum* infection present versus those free of *P. falciparum* infection during the10-year period from 1987 to 1996.

For *P. vivax* malaria, the 131 administrative villages were arbitrarily divided into two groups. ***Those villages with the geometric mean* P. vivax *incidence rates over 2 cases per thousand persons per year over the ten years period (1987-1996) were classified into the high risk group, and those equal to or lower than 2 per thousand persons were classified into the low risk village group.*** The same criteria were also used by the malaria control authority as the threshold to stratify whether or not an administrative village was considered as a high risk village in the county (Guo Hongping, personal communication). Based on those criteria, 53 and 78 administrative villages were classified into the high and low risk groups for *P. vivax* infection, respectively. The mean incidence rate of the high risk group villages was 4.92 per thousand persons per year (with a standard deviation of 2.56, range from 2.06 to 13.24 per thousand persons per year) from 1987 to 1996. The mean incidence rate of the low risk group villages was 0.7 per thousand persons per year (with a standard deviation of 0.68, range from 0 to 1.97 per thousand persons per year) from 1987 to 1996. The difference of the means between two groups was statistically significant using a t-test ($t = -13.88$, $P < 0.001$).

From 1987 to 1996, *P. falciparum* infection had only been reported in 61 administrative villages in Yuanyang. In 70 administrative villages no single *P. falciparum* case had ever been reported during the 10-year period. ***The villages were divided into two groups; group one with no single indigenous falciparum malaria case reported and group two at least one***

indigenous falciparum malaria case reported in the administrative villages from 1987 to 1996.

In order to see the general trend of association between malaria and environmental variables, land use variables in the administrative villages were categorised into three equal sized groups, i.e. tertile groups. Although this will vary from study to study, thus making it difficult to compare findings, it is quite a sensible way of choosing the grouping intervals provided the actual intervals are reported (Clayton & Hill, 1993).

To see general trend, the high versus low risk of *P.vivax* malaria, "with" versus "without" *P. falciparum* malaria administrative village groups were tabulated for the different categories of the environmental variables. The analysis was then to fit one variable at a time into the a logistic regression model to see crude association between the risk of *P. vivax* or *P. falciparum* malaria and each environmental variable. The odds ratio was used to estimate to the strength of the association between malaria and the environmental variable in the model. The test for trend for the risk of malaria at different categories of environmental variables were also carried out as shown in the tables.

The results of crude analyses (Tables 4.7 and 4.8) indicate that the higher the altitude of the villages, the less the risk of malaria. People living in the villages with mean altitude over 1,000 metres had just around one quarter of the risk for vivax malaria and one third of the risk for falciparum malaria of those people living at a mean altitude lower than 1000 metres. Crude analyses also revealed a probable negative correlation between the risk of malaria and the relative amount of dry cropland in the administrative village. The more paddy fields in the villages, the higher risk of malaria for *P. vivax* and *P. falciparum* but they were not statistically significant in the crude modelling. Bush was negatively correlated with vivax malaria. The other land use types do not seem to be the important indicators of malaria.

The association of environmental variables with risk of malaria in the administrative villages might be confounded by other variables. A multivariate logistic regression model for *P. vivax* malaria was fitted by adding one variable at a time, and comparing the fit of the new model with that of the previous model, using the likelihood ratio test. Variables, which improved the fit of the models were retained. Consequently, the altitude, and the relative amounts of paddy and dry cropland were included in the final model. The forest was also included in the final model due to considering its aetiological importance for malaria in the study area. The same set of environmental variables for *P. vivax* malaria were included in the final model of *P. falciparum* malaria, due to the consideration that *P. falciparum* and *P. vivax* have similar aetiological factors and relatively few *P. faclciparum* malaria cases under study, which could result in unstable point estimates during the modelling. The strengths of the association between malaria with other land use variables (the variables did not include in the final model) with malaria were also assessed by adding one variable a time into the final model to see the point estimates. The variables in the final models for both *P. vivax* and *P. falciparum* malaria are shown in the last column of Tables 4.7 and 4.8 as **bold**, respectively.

In the final model, the negative association between the risk of malaria and altitude became stronger as compared with crude analyses. The people living in administrative villages with a mean altitude over 1,000 metres had less than one fifth of the risk of malaria than those living below 1,000 metres. A slightly higher risk was shown in the villages with mean altitude over 1,400 metres than those villages located between 1,000 to 1,400, although 95% confidence intervals of the ORs of two groups overlapped considerably, implying no significant difference. People residing in the higher altitude areas were more likely to migrate to lower altitude areas to work. Therefore, they were likely to get infection in the lowlands. Although we intended to eliminate the imported cases during the stage of the data collection, it is unlikely that we were able to eliminate all the imported cases totally. This misclassification resulted in the falsely higher risk of malaria in the highlands.

The final models indicate a stronger positive association between malaria for both *P. vivax* and *P. falciparum* and the relative amount of paddy in the administrative villages. The initial underestimation of ORs in the crude analysis may be due to the positive correlation between the mean altitude and the relative amount of paddy and forest per administrative village as shown in Table 4.5. The more paddy field uphill on the mountain than in the lower altitude areas might be due to more rainfall and more spring water for irrigation.

The adjusted analyses also reveal a stronger negative association between the risk of vivax malaria and the amount of dry crop field and bush land in the villages (Table 4.7) and a stronger negative association between the risk of falciparum malaria and the relative amounts of dry crop fields and bare rocks (Table 5.8).

The results of the final modelling also suggest a positive correlation of the relative amount of forest and the risk of malaria, although the differences between the two groups are not statistically significant.

Table 4.7. Logistic regression modelling of high risk[1] versus low risk[1] *P.vivax* malaria, altitude and land use variables in the 131 administrative villages, Yuanyang (1987-1996)

Variables[2]	High risk group	Low risk group	Crude odds ratio[3]	Adjusted odds ratio[3,4,5]
Altitude				
<=1000	13(24.53)	7(8.97)	1	**1**
~1400	12 (22.64)	30(38.46)	0.22 (0.07 0.67)	**0.14(0.04 0.57)**
~1600	28 (52.83)	41(52.56)	0.37 (0.13 1.04)	**0.17(0.04 0.69)**
			P=0.03	***P=0.02***
Paddy				
1st tertile(0 - 0.334)	14(26.42)	30(38.46)	1	**1**
2nd tertile(0.335 - 0.463)	17(32.08)	26(33.33)	1.40 (0.58 3.38)	**2.08(0.72 5.97)**
3rd tertile(0.465 - 1.040)	20(41.51)	22(28.21)	2.14 (0.90 5.10)	**7.23(1.89 20.46)**
			P = 0.22	***P = 0.01***
Drycrop				
1st tertile(0 - 0.244)	24(45.28)	20(25.64)	1	**1**
2nd tertile(0.246-.364)	22(39.62)	22(28.21)	0.80 (0.34 1.85)	**0.78(0.30 1.96)**
3rd tertile(0.365-0.792)	36(15.09)	36(46.15)	0.19 (0.07 0.49)	**0.21(0.08 0.61)**
			P =0.002	***P = 0.01***
Tea & fruit land				
1st tertile(0 – 0)	21(39.62)	23(29.49)	1	1
2nd tertile (0-0.082)	18 (33.96)	25(32.05)	0.79 (0.34 1.84)	0.77(0.29 2.08)
3rd tertile(0.083-0.493)	14 (26.42)	30(38.46)	0.51 (0.21 1.22)	0.64(0.23 1.67)
			P = 0.31	*P = 0.63*
Forest				
1st tertile (0-0.332)	18(33.96)	26(33.33)	1	**1**
2nd tertile(0.340-0.620)	13(24.53)	30(38.46)	0.63 (0.26 1.52)	**1.12(0.37 3.41)**
3rd tertile(0.628-1.040)	22 (41.51)	22(28.21)	1.44 (0.62 3.36)	**2.44(0.84 7.83)**
			P = 0.18	***P = 0.15***
Bushes				
1st tertile(0-0.301)	24 (45.28)	20(25.64)	1	1
2nd tertile(0.309-0.510)	12 (22.64)	31(39.74)	0.32 (0.13 0.79)	0.20(0.06 0.64)
3rd tertile(0.520-1.139)	17 (32.08)	27(34.62)	0.52 (0.22 1.23)	0.41(0.12 1.32)
			P = 0.04	*P = 0.02*
Barren land				
1st tertile(0-0.359)	21 (39.62)	23 (29.49)	1	1
2nd tertile(0.374-0.593)	13 (24.53)	30 (38.46)	0.47 (0.20 1.14)	0.93(0.32 2.75)
3rd tertile(0.595-1.114)	19 (35.85)	25 (32.05)	0.83 (0.36 1.93)	1.14(0.31 4.14)
			P = 0.23	*P= 0.95*
Bare rock				
1st tertile (0-0)	17 (32.08)	27 (34.62)	1	1
2nd tertile(0 –0.082)	17 (32.08)	26 (33.33)	1.04 (0.44 2.46)	1.54(0.52 4.54)
3rd tertile(0.092-0.587)	19 (35.85)	25 (32.05)	1.21 (0.52 2.83)	2.18(0.77 6.21)
			P =0.90	*P=0.34*

[1]Those villages with the geometric mean *P. vivax* incidence over 2 per 1000 per year in the ten years were classified into the high risk group, and those equal to or lower than 2 per 1000 annually were allocated to the low risk village group.
[2] Angular transformed proportion of land use variables of administrative villages in parentheses
[3]95% confidence interval in parentheses
[4]Adjusted for altitude, paddy field, dry crop land, forest
[5] Variables in the final model shown as bold

Table 4.8.Logistic regression modelling of with versus without *P.falciparum* malaria, altitude and land use variables in the 131 administrative villages,Yuanyang, (1987-1996)

Variable[1]	With infection	Without infection	Crude odds ratio[2]	Adjusted odds ratio[2,3,4]
Altitude				
<=1000	14(22.95)	6 (8.57)	1	**1**
~1400	15 (24.59)	27(38.57)	0.24(0.08 0.75)	**0.16(0.04 0.63)**
~1600	32 (52.46)	37(52.86)	0.37(0.13 1.08)	**0.18(0.05 0.71)**
			P = 0.049	***P =0.026***
Paddy				
1st tertile(0 - 0.334)	17(27.87)	27(38.57)	1	**1**
2nd tertile(0.335 - 0.463)	19(31.15)	24(34.29)	1.26(0.53 2.96)	**1.62(0.60 4.34)**
3rd tertile(0.465 - 1.040)	25(40.98)	19(27.14)	2.09(0.89 4.89)	**5.62(1.82 17.31)**
			P =0.22	***P= 0.009***
Drycrop				
1st tertile(0 - 0.244)	27(44.26)	17(24.29)	1	**1**
2nd tertile(0.246-.364)	23(37.70)	20(28.57)	0.72(0.31 1.70)	**0.69(0.27 1.75)**
3rd tertile(0.365-0.792)	11(18.03)	33(47.14)	0.21(0.08 0.52)	**0.25(0.09 0.68)**
			P = 0.002	***P = 0.02***
Tea & fruit land				
1st tertile(0 – 0)	25(40.98)	19(27.14)	1	1
2nd tertile (0-0.082)	19(31.15)	24(34.29)	0.60(0.26 1.40)	0.53(0.20 1.39)
3rd tertile(0.083-0.493)	17(27.87)	27(38.57)	0.48(0.20 1.12)	0.55(0.21 1.42)
			P = 0.22	*P = 0.34*
Forest				
1st tertile (0-0.332)	19(31.15)	25(35.71)	1	**1**
2nd tertile(0.340-0.620)	19(31.15)	24(34.29)	1.04(0.45 2.43)	**0.98(0.70 5.66)**
3rd tertile(0.628-1.040)	23(37.70)	21(30.00)	1.44(0.62 3.34)	**2.45(0.85 7.04)**
			P = 0.65	***P= 0.24***
Bushes				
1st tertile(0-0.301)	24(39.34)	20(28.57)	1	1
2nd tertile(0.309-0.510)	18(29.51)	25(35.71)	0.60(0.26 1.40)	0.55(0.20 1.50)
3rd tertile(0.520-1.139)	19(31.15)	25(35.71)	0.63(0.27 1.47)	0.59(0.20 1.77)
			P= 0.43	*P=0.47*
Barren land				
1st tertile(0-0.359)	23(37.70)	21(30.00)	1	1
2nd tertile(0.374-0.593)	18(29.51)	25(35.71)	0.66(0.28 1.53)	1.38(0.49 3.90)
3rd tertile(0.595-1.114)	20(32.79)	24(34.29)	0.76(0.33 1.76)	1.07(0.31 3.72)
			P =0.61	*P= 0.79*
Bare rock				
1st tertile (0-0)	20(32.79)	24(34.29)	1	1
2nd tertile(0 –0.082)	16(26.23)	27(38.57)	0.71(0.30 1.68)	0.81(0.30 2.24)
3rd tertile(0.092-0.587)	25(40.98)	19(27.14)	1.58(0.68 3.66)	2.65(0.97 7.24)
			P=0.19	*P=0.05*

[1]Angular transformed proportion of land use variables of administrative villages in parentheses
[2]95% confidence interval in parentheses
[3]Adjusted for altitude, paddy field, dry crop land, forest
[4]Variables in the final model shown as bold

The association between the soils and malaria for *P. vivax* and *P. falciparum* was also assessed by logistic regression modelling as shown in Tables 4.9 and 4.10. Crude analyses indicate that the relative amounts of brick red soil, red soil, yellow soil and yellow brown soil were significantly negatively associated with the risk of vivax malaria in the village. Crude analyses also suggested that the relative amount of brick red soil and yellow brown soil was significantly negatively correlated with the risk of falciparum malaria in the village. The dry red soil was positively associated with *P. falciparum* malaria infection in the villages. Nevertheless, the association between soil types and malaria might be confounded by other potential confounding variables. Therefore, the associations were further assessed by adding one soil type variable a time to the final models of logistic regression to see whether the associations still existed. As shown in Figures 4.13 and 4.14, paddy fields in land use map and paddy soil basically overlapped, occupying the same area spatially. The original classification of soils for paddy soil might be based on the land use type, paddy field. Therefore, in the present analysis, we didn't adjust for the paddy field variable, as it is synonymous with the paddy soil variable. As expected, the results of the analysis indicate that the paddy soil was positively correlated with the risk of malaria (Tables 4.9 and 4.10), giving similar results to those described for the paddy field (Tables 4.7 and 4.8). The apparent association of the relative amount of red soil, yellow soil and yellow brown soil and risk of vivax malaria disappear after adjusting for confounding variables although brick red soil is still at the borderline. The positive association of dry red soil and risk of *P. falciparum* malaria infection remained the same after adjusting for other confounding variables in the final model. The apparent association of the relative amount of yellow brown soil and brick red soil and risk of *P. falciparum* malaria disappear after adjusting for confounding variables. The results suggest that apparent associations between most soils and the risk of malaria for both *P. vivax* and *P. falciparum* were due to a confounding effect of altitude and other variables in the administrative villages.

The paddy rice field provided the vital environment for mosquito breeding sites in the malaria transmission season. Therefore, it is paddy rice field that increased malaria risk in villages. Nevertheless, the original soil classification is hardly associated with mosquito ecology except for paddy soil. A soil classification based on their water retaining capacity would allow more meaningful analysis in the future studies.

Table 4.9. Logistic regression modelling of high risk[1] versus low risk[1] *P.vivax* malaria and soils in the 131 administrative villages, Yuanyang (1987-1996)

Variables[2]	High risk group	Low risk group	Crude odds ratio[3]	Adjusted odds ratio[3,4]
Paddy soil				
1st tertile(0-0.338)	17(32.08)	27(34.62)	1	1
2nd tertile(0.340-0.473)	15(28.30)	28(35.90)	0.85(0.36 2.04)	1.48(0.52 4.24)
3rd tertile(0.474-1.135)	21(39.62)	23(29.49)	1.45(0.62 3.38)	4.40(1.37 14.18)
			P=0.46	*P=0.03*
Brick red soil				
1st tertile(0-0)	22(41.51)	22(28.21)	1	1
2nd tertile(0-0)	22(41.51)	21(26.92)	1.05(0.45 2.43)	0.77(0.27 2.18)
3rd tertile(0-0.648)	9(16.98)	35(44.87)	0.26(0.10 0.66)	0.30(0.10 0.88)
			P=0.006	*P=0.07*
Dark red soil				
1st tertile(0-0.320)	19(35.85)	25(32.05)	1	1
2nd tertile(0.328-0.585)	15(28.30)	28(35.90)	0.70(0.30 1.67)	0.43(0.13 1.43)
3rd tertile(0.590-0.957)	19(35.85)	25(32.05)	1.00(0.43 2.32)	0.60(0.16 2.18)
			P=0.66	*P=0.38*
Red soil				
1st tertile(0-0.348)	25(47.17)	19(24.36)	1	1
2nd tertile(0.358-0.542)	13 (24.53)	30(38.46)	0.33(0.14 0.80)	0.39(0.13 1.14)
3rd tertile(0.545-0.877)	15 (28.30)	29(37.18)	0.39(0.17 0.93)	0.42(0.14 1.21)
			P=0.03	*P=0.16*
Yellow soil				
1st tertile(0-0.315)	24(45.28)	20(25.64)	1	1
2nd tertile(0.316-0.522)	10(18.87)	33(42.31)	0.25(0.10 0.64)	0.44(0.12 1.61)
3rd tertile(0.528-0.868)	19(35.85)	25(32.05)	0.63(0.27 1.47)	1.13(0.23 5.44)
			P=0.013	*P=0.19*
Dry red soil				
1st tertile(0-0)	15(28.30)	29(37.18)	1	1
2nd tertile(0-0)	15(28.30)	28(35.90)	1.04(0.43 2.51)	1.54(0.54 4.35)
3rd tertile(0-1.28)	23(43.40)	21(26.92)	2.12(0.90 5.00)	1.47(0.43 5.04)
			P=0.15	*P=0.69*
Yellow brown soil				
1st tertile(0-0)	23(43.40)	21(26.92)	1	1
2nd tertile(0-0.218)	18(33.96)	25(32.05)	0.66(0.28 1.53)	0.65(0.24 1.73)
3rd tertile(0.257-1.186)	12(22.64)	32(41.03)	0.34(0.14 0.83)	0.69(0.19 2.55)
			P=0.006	*P=0.67*
Brown soil				
1st tertile(0-0)	20(37.74)	24 (30.77)	1	1
2nd tertile(0-0)	18(33.96)	25(32.05)	0.86(0.37 2.02)	0.97(0.36 2.66)
3rd tertile(0-1.038)	15(28.30)	29 (37.18)	0.62(0.26 1.47)	0.48(0.17 1.35)
			P=0.54	*P=0.30*
Purple soil				
1st tertile(0-0)	14(26.42)	30(38.46)	1	1
2nd tertile(0-0)	20(37.74)	23(29.49)	1.86(0.78 4.46)	1.47(0.54 4.03)
3rd tertile(0-0.250)	19(35.85)	25(32.05)	1.63(0.68 3.89)	1.72(0.62 4.73)
			P=0.34	*P=0.56*

[1]Those villages with the geometric mean *P. vivax* incidence over 2 per thousand per year in the past ten years were classified into the higher risk group, and those equal to or lower than 2 per thousand annually were allocated to the lower risk village group.
[2] Angular transformed proportion of land use variables of administrative villages in parentheses
[3]95% confidence interval in parentheses
[4]Adjusted for altitude, paddy field, dry crop land, forest

Table 4.10.Logistic regression modelling of with versus without *P.falciparum* malaria and soils in the 131 administrative villages,Yuanyang, Yunnan (1987-1996)

Soils[1]	With infection	Without infection	Crude odds ratio[2]	Adjusted odds ratio[2,3]
Paddy soil				
1st tertile(0-0.338)	19(31.15)	25(35.71)	1	1
2nd tertile(0.340-0.473)	17(27.87)	26(37.14)	0.86(0.37 2.02)	1.64(0.59 4.60)
3rd tertile(0.474-1.135)	25(40.98)	19(27.14)	1.73(0.74 4.02)	6.96(2.08 23.26)
			P=0.24	*P=0.008*
Brick red soil				
1st tertile(0-0)	23(37.70)	21(30.00)	1	1
2nd tertile(0-0)	25(40.98)	18(25.71)	1.27(0.54 2.96)	1.02(0.38 2.70)
3rd tertile(0-0.648)	13(21.31)	31(44.29)	0.38(0.16 0.92)	0.40(0.15 1.10)
			P=0.02	*P=0.12*
Dark red soil				
1st tertile(0-0.320)	20(32.79)	24(34.29)	1	1
2nd tertile(0.328-0.585)	18(29.51)	25(35.71)	0.86(0.37 2.02)	0.55(0.18 1.68)
3rd tertile(0.590-0.957)	23(37.70)	21(30.00)	1.31(0.57 3.04)	0.85(0.26 2.80)
			P=0.61	*P=0.51*
Red soil				
1st tertile(0-0.348)	25(40.98)	19(27.14)	1	1
2nd tertile(0.358-0.542)	19(31.15)	24(34.29)	0.60(0.26 1.40)	0.84(0.31 2.30)
3rd tertile(0.545-0.877)	17(27.87)	27(38.57)	0.48(0.20 1.12)	0.53(0.19 1.45)
			P=0.22	*P=0.43*
Yellow soil				
1st tertile(0-0.315)	25(40.98)	19(27.14)	1	1
2nd tertile(0.316-0.522)	16(26.23)	27(38.57)	0.45(0.19 1.06)	1.05(0.32 3.50)
3rd tertile(0.528-0.868)	20(32.79)	24(34.29)	0.63(0.27 1.47)	1.55(0.33 7.28)
			P=0.19	*P=0.78*
Dry red soil				
1st tertile(0-0)	18(29.51)	26(37.14)	1	1
2nd tertile(0-0)	13(21.31)	30(42.86)	0.63(0.26 1.52)	0.75(0.28 2.01)
3rd tertile(0-1.28)	30(49.18)	14(20.00)	3.10(1.29 7.42)	3.76(1.10 12.84)
			P=0.002	*P=0.03*
Yellow brown soil				
1st tertile(0-0)	26(42.62)	18 (25.71)	1	1
2nd tertile(0-0.218)	21(34.43)	22 (31.43)	0.66(0.28 1.54)	0.66(0.26 1.69)
3rd tertile(0.257-1.186)	14(22.95)	30 (42.86)	0.32(0.13 0.77)	0.62(0.18 2.12)
			P=0.04	*P=0.63*
Brown soil				
1st tertile(0-0)	21(34.43)	23(32.86)	1	1
2nd tertile(0-0)	24(39.34)	19(27.14)	1.38(0.59 3.22)	1.77(0.66 4.71)
3rd tertile(0-1.038)	16(26.23)	28(40.00)	0.63(0.27 1.47)	0.53(0.20 1.41)
			P=0.19	*P=0.07*
Purple soil				
1st tertile(0-0)	20(32.79)	24(34.29)	1	1
2nd tertile(0-0)	21(34.43)	22(31.43)	1.15(0.49 2.66)	0.90(0.35 2.32)
3rd tertile(0-0.250)	20(32.79)	24(34.29)	1.00(0.43 2.31)	1.01(0.39 2.62)
			P=0.94	*P=0.97*

[1]Angular transformed proportion of land use variables of administrative villages in parentheses
[2]95% confidence interval in parentheses
[3]Adjusted for altitude, paddy field, dry crop land, forest

4.4.4 Multilevel cluster analysis

4.4.4.1 Malaria clustering within administrative township

The preliminary analysis described above was based on the assumption that the malaria risks are independent among the administrative villages. But if we visually check the spatial distribution of malaria in the Yuanyang, we find that malaria tends to be spatially auto-correlated (grouped spatially) as shown in Figures 4.7 and 4.8. The risks of malaria in adjacent administrative villages are much more similar than are the risks in far distant villages. Spatial autocorrelation of malaria among administrative villages is likely to be due to ecological factor aggregation, which favours similar malaria transmission among the clustering villages. The ecological situations could be defined as altitude, paddy, forest etc., which we could model. On the other hand, the spatial autocorrelation could be due to variables which we did not measure or couldnot be measured such as the socio-economic situation among the villages and the performances of data collection in different towns and villages (Ecob, 1996).

The unit of malaria surveillance and reporting for the villages was the township hospital in Yuanyang County as shown in Figure 4.5. The quality of data collected in a township was highly dependent on how well the staff performed in the hospital disease control group in the township hospitals of each administrative township. Therefore, the number of cases detected and reported is likely to be determined Partly by the performance of the team of the malaria control group in the township hospital. The performance of village doctors in routine surveillance is also likely to depend on peer pressure within the township since the village doctors would meet once a month with the malaria control group in the township hospital. Malaria is a vector-borne communicable disease, therefore the probability of people infecting each other within a cluster is very high. Hence the spatial autocorrelation among the villages was likely to be partly due to the malaria case reporting system of the township and transmission of the disease within the community.

Table 4.11 shows the incidence rates of *P. vivax* and *P. falciparum* malaria in the townships in Yuangyang over the10-year period from 1987 to 1996. As shown in the table, there was great variation in the number of malaria cases reported among townships over the 10-year period. The lowest vivax malaria incidence rate for an administrative township was 0.07 per 1000 persons, and the highest one was 8.17 per 1000 persons. Two townships had not a single case report of *P. falciparum* infection over the 10-year period. The *highest P. falciparum* infection risk of a township was 2.42 per1000 persons over the 10-year period. Malaria variation among townships might be partly due to reasons described above.

Table 4.11 *P. vivax* and *P. falciparum* malaria distribution in the15 Administrative Townships in Yuanyang County, Yunnan (1987-1996)

Town.	No.Pop.	No.vivax	Incidence/1000[1]	No. falcip.	Incidence/1000[1]
Xinjie	258,712	1,569	6.07(5.77, 6.37)	81	0.31(0.25, 0.39)
Nansa	114,100	932	8.17(7.66, 8.71)	46	0.40(0.30, 0.54)
Salatuo	192,278	92	0.48(0.39, 0.59)	21	0.11(0.07, 0.17)
Shenchen	321,928	1,258	3.91(3.70, 4.13)	10	0.03(0.02, 0.06)
Nujuezai	289,159	982	3.40(3.19, 3.62)	699	2.42(2.25, 2.60)
Gannian	174,114	62	0.36(0.28, 0.46)	1	0.05(0.001, 0.04)
Sanxinchem	196,761	113	0.57(0.48, 0.69)	7	0.04(0.02, 0.08)
Xiaoxinjie	218,115	365	1.67(1.51, 1.85)	2	0.01(0.002, 0.04)
Feng Chun Ling	308,788	2,352	7.62(7.32, 7.93)	421	1.36(1.24, 1.50)
Daping	192,396	843	4.38(4.10, 4.69)	3	0.02(0.005, 0.05)
Penzihua	154,256	551	3.57(3.29, 3.88)	4	0.03(0.01, 0.07)
Fanmaolin	158,912	432	2.72(2.47, 2.99)	6	0.04(0.02, 0.08)
Huanchaolin	284,140	20	0.07(0.05, 0.11)	0	
Ezha	181,851	43	0.24(0.17, 0.32)	0	
Majie	252,475	1,617	6.41(6.10, 6.72)	80	0.32(0.26, 0.39)

[1]95% confidence intervals in parentheses

4.4.4.2 Multilevel model

A multilevel model is a tool to analyse data with a hierarchical structure and allows us to take into account cluster correlation (Longford, 1993; Kreft & Leeuw, 1998) or so called "*group effects*" (Goldstein, 1995). Ignoring the *group effects* will result in falsely precise estimate (Jones & Duncan, 1996). In the present study we have two levels, administrative village and township. Here, we use simple linear regression with two levels as an example to understand multilevel modeling in the present study. Suppose Y_{ij} is the response for administrative village i within township j. The response under study may depend upon a group of covariates X_{ij}, with corresponding regression co-efficients β:

$$Y_{ij} = \beta_o + \beta X_{ij} + \epsilon_{ij}$$

Where ϵ_{ij} is the normal model residual. However, one individual administrative village might be more similar to another from the same township and hence the intercept might vary between townships. Therefore a two-level model is:

$$Y_{ij} = \beta_o + \beta X_{ij} + \mu_j + \epsilon_{ij}$$

Where the response is the sum of a fixed effect, $\beta_o + \beta X_{ij}$, and a random effect $\mu_j + \epsilon_{ij}$.. μ_j is a random effect with variance σ^2_μ reflecting the variation between townships. ϵ_{ij} is a random effect with variance of σ^2_ϵ reflecting the variation between the administrative villages. This is the basic structure of a two-level linear model.

In the present study, logistic regression was used to model the odds of being a high risk or low risk vivax malaria administrative village and the odds of being in villages with and without falciparum malaria infection over the past 10 year period with their landscape and environmental variables. In this case, let Y_{ij} be a binary response (high versus low risk vivax malaria groups, with versus without falciparum malaria groups) for individual administrative village *i* in township *j*. The logistic regression model can be written as (Goldstein, 1995):

$$Y_{ij} = \pi_{ij} + \epsilon_{ij}$$

Where

$$\mathrm{Log}\{\pi_{ij}/(1-\pi_{ij})\} = a + \beta X_{ij} + u_j$$

Therefore, the response Y_{ij} is the sum of a fixed part, $a+\beta X_{ij}$, and random part, $\mu_j + \epsilon_{ij}$. a is the vector of parameters corresponding to vector of covariates. μ_j is a random effect with variance of σ^2_μ reflecting the variation between townships. The level 1 random part ϵ_{ij} is defined by a dummy variable equal to the square root of the expected count and variance constrained to 1 when assuming a logistic distribution (Goldstein, 1995 & Goldstein *et al.*, 1998).

The model was estimated by using *MlwiN* (Goldstein *et al.*, 1998). The estimate procedure is iterative with second-order penalized quasi-likelihood (Goldstein and Rasbash, 1996). Parsimonious models included all the variables included in the ordinary logistic regression models for *P. vivax* and *P. falciparum* malaria as described in section 4.4.3 in this Chapter.

4.4.4.3 The results of multilevel logistic regression modeling

The risk factors for vivax and falciparum malaria (high versus low risk administrative village groups, administrative village with versus without *P. falciparum* malaria) were analyzed by fitting a multilevel logistic regression model. Based on the results of the ordinary logistic regression analyses, the variables of the mean altitude, the relative amount of paddy, forest and dry cropland in the administrative villages were included into the final multilevel model.

The antilog values of the coefficients are the estimated odds ratio (Armitage and Berry, 1987). The final multilevel logistic regression models for *P. vivax* and *P. falciparum* malaria are shown as **bold** in the last column of Tables 4.12 and 4.13, respectively. The other variables were also analyzed by adding one variable a time into the final model to appraise their aetiological importance on malaria which is also shown in the last column. The results of the final models of the ordinary logistic regression for *P. vivax* and *P. falciparum* malaria are also shown in the second column of the tables for comparison.

The final multilevel models for *P. vivax* and *P. falciparum* malaria and mean altitude are shown in Tables 4.12 and 4.13, respectively. In the multilevel logistic regression model, the negative association between altitude and risk of vivax malaria is stronger, particularly in the middle altitude administrative village category. The results suggested a strong *group effect* with altitude. It might be due to more vivax malaria cases being reported in the villages located at middle altitudes, or *P. vivax* transmitted among the villages in the middle altitude area. The negative association between risk of falciparum malaria and mean altitude as shown in Table 4.13, however, became weaker in the multilevel model, particularly in the high altitude administrative village categories on removing the *group effect* in the multilevel model.

A strong positive association between the risk of vivax malaria and the amount of paddy within an administrative village was shown in the final multilevel logistic regression model. But the association between the two variables is weaker as compared the results with that in non-spatial ordinary logistic regression analysis, suggesting a stronger spatial autocorrelation in the townships with villages containing more paddy fields. It might be that more attention had been paid to areas with more paddy in routine surveillance systems, or malaria transmission occurred within the villages with more paddy fields or other characteristics of the villages with more paddy fields that we did not collect nor included them in the model. The negative association between falciparum malaria and paddy remained roughly the same in the final multilevel model as compared with that in non-spatial ordinary logistic regression analysis as shown in Table 4.13.

The appearance of a negative association between dry cropland and bush and vivax malaria in the administrative villages, demonstrated in the ordinary logistic regression model, disappears in the final multilevel model. Similarly, the appearance of a negative association between the relative amount of dry cropland and the risk of *P. faciparum* infection also disappeared in the final model for *P. falciparum* malaria as did the appearance of a positive association between bare rock and risk of *P. falciparum* infection. The results suggest that the appearance of negative associations of those variables with malaria was largely due to spatial autocorrelation within administrative villages. It might be because less malaria cases were reported in this area in the routine malaria surveillance systems.

The association between forest and the risk of vivax and falciparum malaria in the final multilevel logistic regression model was shown in Tables 4.12 and 4.13. There was a 2-fold higher risk in 2nd and 3rd tertiles of forest for vivax malaria as compared with that 1st tertile of forest, suggesting a possible positive association between two variables. But this is not statistically significant.

A similar method was also used to analyze the effect of other land use types by adding one variable at a time to the final multilevel logistic regression models such as tea and tropical fruit land, barren land and bare rock. There was no evidence of the association between risk of malaria in administrative villages and the land use variables of tea and fruit land, barren land and bare rock by multilevel modeling as shown in the last columns of Tables 4.12 and 4.13.

The random parts of multilevel logistic regression models for *P. vivax* and *P. falciparum* malaria were shown in the last rows of Tables 4.12 and 4.13, respectively. The random parts of township level variance are at borderline significance for the *P. vivax* malaria model but significant *for P. falciparum* malaria model, suggesting a strong spatial autocorrelation (*group effects*) of malaria within township. The variance could not be explained by variables included in the final model, presumably due to the performance of township hospital staff's work, or other characteristics of the township which we did not collect and model in the multilevel model.

Table 4.12. Multilevel logistic regression modelling of high risk[1] versus low risk[1] *P.vivax* malaria, altitude and land use variables in the 131 administrative villages, Yuanyang (1987-1996)

Variable[2]	Adjusted odds ratio[3]	Crude odds ratio[4]	Adjusted odds ratios[5,6]
Altitude			
<=1000	**1**	1	**1**
~1400	**0.14(0.04 0.57)**	0.08(0.01 0.69)	**0.07(0.08 0.66)**
~1600	**0.17(0.04 0.69)**	0.23(0.03 1.69)	**0.14(0.01 1.24)**
Paddy			
1st tertile(0 - 0.334)	**1**	1	**1**
2nd tertile(0.335 - 0.463)	**2.08(0.72 5.97)**	1.75(0.43 7.21)	**2.30(0.34 6.29)**
3rd tertile(0.465 - 1.040)	**7.23(1.89 20.46)**	3.23(0.67 15.55)	**5.73(0.96 24.61)**
Drycrop			
1st tertile(0 - 0.244)	**1**	1	**1**
2nd tertile(0.246-.364)	**0.78(0.30 1.96)**	1.36(0.40 4.64)	**1.27(0.34 4.66)**
3rd tertile(0.365-0.792)	**0.21(0.08 0.61)**	1.90(0.44 8.20)	**0.64(0.14 2.81)**
Tea & fruit land			
1st tertile(0 – 0)	1	1	1
2nd tertile (0-0.082)	0.77(0.29 2.08)	0.79(0.24 2.72)	0.75(0.20 2.77)
3rd tertile(0.083-0.493)	0.64(0.23 1.67)	0.50(0.12 2.10)	0.42(0.09 10.81)
Forest			
1st tertile (0-0.332)	**1**	1	**1**
2nd tertile(0.340-0.620)	**1.12(0.37 3.41)**	1.10(0.28 4.32)	**2.30(0.47 11.35)**
3rd tertile(0.628-1.040)	**2.44(0.84 7.83)**	0.89(0.25 3.19)	**2.14(0.31 10.64)**
Bushes			
1st tertile(0-0.301)	1	1	1
2nd tertile(0.309-0.510)	0.20(0.06 0.64)	0.38(0.09 1.54)	0.36(0.07 1.88)
3rd tertile(0.520-1.139)	0.41(0.12 1.32)	1.17(0.27 5.06)	1.44(0.22 9.17)
Barren land			
1st tertile(0-0.359)	1	1	1
2nd tertile(0.374-0.593)	0.93(0.32 2.75)	0.81(0.21 3.07)	1.14(0.25 5.20)
3rd tertile(0.595-1.114)	1.14(0.31 4.14)	1.05(0.28 3.88)	1.59(0.22 11.48)
Bare rock			
1st tertile (0-0)	1	1	1
2nd tertile(0 –0.082)	1.54(0.52 4.54)	0.98(0.28 3.48)	1.29(0.28 4.41)
3rd tertile(0.092-0.587)	2.18(0.77 6.21	0.47(0.11 1.90)	0.84(0.19 3.75)
Random effect			
Township level variance(SE)			**3.84(2.02)**
Village level variance (SE)			**1**

[1]Those villages with the geometric mean *P. vivax* incidence over 2 per thousand per year in the past ten years were classified into the higher risk group, and those equal to or lower than 2 per thousand annually were classified into the lower risk village group.
[2]Angular transformed proportion of land use variables of administrative villages in parentheses
[3]Odds ratio with 95% confidence interval in parentheses in ordinary logistic regressuion
[4]Crude odd ratios with 95% confidence interval in parentheses in multilevel logistic regressuion
[5]Adjusted Odd ratios for altitude, paddy field, dry crop land, forest with 95% confidence interval in parentheses in multilevel logistic regressuion
[6] Variables in the final multilevel model shown as bold

Table 4.13. Multilevel logistic regression modelling of with versus without *P.falciparum* malaria[1]: altitude and land use variables in the 131 administrative villages,Yuanyang,(1987-1996)

Variables[2]	Adjusted odds ratio[3]	Crude odds ratio[4]	Adjusted odds ratio[5,6]
Altitude			
<=1000	**1**	1	**1**
~1400	**0.16(0.04 0.63)**	0.39(0.07 2.32)	**0.24(0.04 1.81)**
~1600	**0.18(0.05 0.71)**	0.84(0.16 4.36)	**0.46(0.07 3.05)**
Paddy			
1st tertile(0 - 0.334)	**1**	1	**1**
2nd tertile(0.335 - 0.463)	**1.62(0.60 4.34)**	2.12(0.51 8.91)	**2.13(0.49 9.16)**
3rd tertile(0.465 - 1.040)	**5.62(1.82 17.31)**	4.43(0.96 20.51)	**6.20(1.13 34.10)**
Drycrop			
1st tertile(0 - 0.244)	**1**	1	**1**
2nd tertile(0.246-.364)	**0.69(0.27 1.75)**	1.20(0.34 4.21)	**1.07(0.29 3.94)**
3rd tertile(0.365-0.792)	**0.25(0.09 0.68)**	0.61(0.15 2.44)	**0.61(0.14 2.56)**
Tea & fruit land			
1st tertile(0 – 0)	1	1	1
2nd tertile (0-0.082)	0.53(0.20 1.39)	0.54(0.17 1.72)	0.38(0.12 1.33)
3rd tertile(0.083-0.493)	0.55(0.21 1.42)	0.49(0.13 1.88)	0.44(0.19 1.73)
Forest			
1st tertile (0-0.332)	**1**	1	**1**
2nd tertile(0.340-0.620)	**0.98(0.70 5.66)**	1.97(0.55 7.11)	**2.85(0.67 12.11)**
3rd tertile(0.628-1.040)	**2.45(0.85 7.04)**	0.87(0.24 3.13)	**1.33(0.32 5.67)**
Bushes			
1st tertile(0-0.301)	1	1	1
2nd tertile(0.309-0.510)	0.55(0.20 1.50)	0.82(0.24 3.16)	1.00(0.23 4.21)
3rd tertile(0.520-1.139)	0.59(0.20 1.77	1.16(0.28 4.76)	1.98(0.61 11.42)
Barren land			
1st tertile(0-0.359)	1	1	1
2nd tertile(0.374-0.593)	1.38(0.49 3.90)	0.97(0.27 3.48)	1.57(0.37 6.68)
3rd tertile(0.595-1.114)	1.07(0.31 3.72)	0.76(0.20 2.91)	1.49(0.20 11.09)
Bare rock			
1st tertile (0-0)	1	1	1
2nd tertile(0 –0.082)	0.81(0.30 2.24)	0.81(0.22 2.96)	0.91(0.22 3.66)
3rd tertile(0.092-0.587)	2.65(0.97 7.24)	1.57(0.40 6.21)	2.59(0.56 11.89)
Random effect			
Township level variance(SE)			**3.69(1.83)**
Village level variance (SE)			**1**

[1]Those villages with no single *P. falciparum* malaria infection were compared with those villages with at least one case of *P. falciparum* malaria infection.
[2]Angular transformed proportion of land use variables of administrative villages in parentheses
[3]Odds ratio with 95% confidence interval in parentheses in ordinary logistic regressuion
[4]Crude odd ratios with 95% confidence interval in parentheses in the multilevel logistic regressuion
[5]Adjusted Odd ratios for altitude, paddy field, dry crop land, forest with 95% confidence interval in parentheses in multilevel logistic regression.
[6] Variables in the final multilevel model shown as bold

The association between soils and their risks of for both *P. vivax* and *P. falciparum* malaria was also assessed by adding one soil type variable at a time into the final multilevel logistic regression models for analyses. The results of the analysis were shown in Tables 4.14 and 4.15. We did not adjust for paddy variable, when we analysed the effect of paddy soil on the risk of malaria, as already discussed in section 4.4.3 of this Chapter. The results of analysis indicated that the amount of paddy soil was positively correlated with the risk of malaria in the administrative villages in Yuanyang. The result was consistent with that of the analysis of land use variables with the risk of vivax malaria. Similarly, a positive association was identified between risk of falciparum malaria and paddy soil as shown in Table 4.15. The apparent positive association between relative amounts of dry red soil and risk of vivax malaria disappeared after adjusting for confounding variables as well as spatial autocorrelation within the townships. It suggests that apparent associations of the soils and the risk of malaria are mostly due to the confounding effect of altitude and other confounding variables in the administrative villages

Table 4.14. Multilevel logistic regression modelling for high versus low risk[1] *P.vivax* malaria and soils in the 131 administrative villages,Yuanyang(1987-1996)

Soils[2]	Adjusted odds ratio[3]	Crude odds ratio[4]	Adjusted odds ratio[5]
Paddy soil			
1st tertile(0-0.338)	1	1	1
2nd tertile(0.340-0.473)	1.48(0.52 4.24)	0.99(0.22 4.42)	1.47(0.34 6.40)
3rd tertile(0.474-1.135)	4.40(1.37 14.18)	1.75(0.33 9.18)	4.92(0.77 31.45)
Brick red soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	0.77(0.27 2.18)	0.89(0.25 3.20)	0.79(0.20 3.15)
3rd tertile(0-0.648)	0.30(0.10 0.88)	0.31(0.08 1.21)	0.37(0.09 1.52)
Dark red soil			
1st tertile(0-0.320)	1	1	1
2nd tertile(0.328-0.585)	0.43(0.13 1.43)	0.40(0.11 1.48)	0.18(0.03 1.06)
3rd tertile(0.590-0.957)	0.60(0.16 2.18)	0.56(0.16 1.96)	0.21(0.03 1.69)
Red soil			
1st tertile(0-0.348)	1	1	1
2nd tertile(0.358-0.542)	0.39(0.13 1.14)	0.27(0.07 1.10)	0.36(0.09 1.50)
3rd tertile(0.545-0.877)	0.42(0.14 1.21)	0.36(0.09 1.34)	0.36(0.09 1.45)
Yellow soil			
1st tertile(0-0.315)	1	1	1
2nd tertile(0.316-0.522)	0.44(0.12 1.61)	0.29(0.08 1.10)	0.28(0.05 1.55)
3rd tertile(0.528-0.868)	1.13(0.23 5.44)	0.54(0.15 1.93)	0.54(0.07 4.44)
Dry red soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	1.54(0.54 4.35)	1.65(0.46 5.89)	1.43(0.78 5.05)
3rd tertile(0-1.28)	1.47(0.43 5.04)	1.35(0.34 5.26)	1.66(0.29 9.57)
Yellow brown soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0.218)	0.65(0.24 1.73)	0.33(0.09 1.26)	0.40(0.10 1.66)
3rd tertile(0.257-1.186)	0.69(0.19 2.55)	0.40(0.10 1.63)	1.08(0.16 7.14)
Brown soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	0.97(0.36 2.66)	0.81(0.20 3.30)	0.95(0.24 3.68)
3rd tertile(0-1.038)	0.48(0.17 1.35)	0.47(0.13 1.75)	0.42(0.11 1.67)
Purple soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	1.47(0.54 4.03)	1.20(0.53 5.17)	1.37(0.75 5.07)
3rd tertile(0-0.250)	1.72(0.62 4.73)	2.09(0.56 7.80)	1.26(0.61 8.40)

[1]Those villages with the geometric mean *P. vivax* incidence over 2 per thousand per year in the past ten years were classified into the higher risk group, and those equal to or lower than 2 per thousand annually were allocated into the lower risk village group.
[2]Angular transformed proportion of soil variables of administrative villages in parentheses
[3]Odds ratio with 95% confidence interval in parentheses in ordinary logistic regressuion
[4]Crude odd ratios with 95% confidence interval in parentheses in multilevel logistic regressuion
[5]Adjusted Odd ratios for altitude, paddy field, dry crop land, forest with 95% confidence interval in parentheses in multilevel logistic regressuion

Table 4.15. Multilevel logistic regression modelling of with, versus without, *P.falciparum* malaria[1] and soil variables in the 131 administrative villages,Yuanyang, (1987-1996)

Soils[2]	Adjusted odds ratio[3]	Crude odds ratio[4]	Adjusted odds ratio[5]
Paddy soil			
1st tertile(0-0.338)	1	1	1
2nd tertile(0.340-0.473)	1.64(0.59 4.60)	1.28(0.34 4.78)	1.49(0.36 6.12)
3rd tertile(0.474-1.135)	6.96(2.08 23.26)	2.87(0.79 10.41)	4.12(0.91 18.65)
Brick red soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	1.02(0.38 2.70)	1.70(0.47 6.16)	1.64(0.43 6.31)
3rd tertile(0-0.648)	0.40(0.15 1.10)	1.78(0.48 6.50)	2.18(0.49 9.63)
Dark red soil			
1st tertile(0-0.320)	1	1	1
2nd tertile(0.328-0.585)	0.55(0.18 1.68)	0.66(0.19 2.31)	0.43(0.10 1.88)
3rd tertile(0.590-0.957)	0.85(0.26 2.80)	0.88(0.27 2.82)	0.74(0.15 3.69)
Red soil			
1st tertile(0-0.348)	1	1	1
2nd tertile(0.358-0.542)	0.84(0.31 2.30)	1.29(0.39 4.29)	1.27(0.36 4.50)
3rd tertile(0.545-0.877)	0.53(0.19 1.45)	0.97(0.27 3.51)	0.92(0.23 3.62)
Yellow soil			
1st tertile(0-0.315)	1	1	1
2nd tertile(0.316-0.522)	1.05(0.32 3.50)	1.65(0.72 5.80)	1.24(0.45 7.18)
3rd tertile(0.528-0.868)	1.55(0.33 7.28	0.80(0.22 2.90)	0.59(0.08 4.15)
Dry red soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	0.75(0.28 2.01)	0.77(0.21 2.78)	0.86(0.23 3.15)
3rd tertile(0-1.28)	3.76(1.10 12.84	2.21(0.75 6.50)	2.66(0.65 9.07)
Yellow brown soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0.218)	0.66(0.26 1.69)	0.71(0.20 2.58)	0.80(0.21 3.05)
3rd tertile(0.257-1.186)	0.62(0.18 2.12	0.74(0.22 2.54)	1.28(0.25 6.67)
Brown soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	1.77(0.66 4.71)	1.38(0.43 4.41)	1.86(0.37 9.39)
3rd tertile(0-1.038)	0.53(0.20 1.41	0.36(0.09 1.42)	0.42(0.09 1.97)
Purple soil			
1st tertile(0-0)	1	1	1
2nd tertile(0-0)	0.90(0.35 2.32)	0.27(0.07 1.08)	0.25(0.06 1.04)
3rd tertile(0-0.250)	1.01(0.39 2.62)	0.36(0.11 1.22)	0.42(0.12 1.53)

[1]Those villages with the geometric mean *P. vivax* incidence over 2 per thousand per year in the past ten years were classified into the higher risk group, and those equal to or lower than 2 per thousand annually were classified into the lower risk village group.
[2]Angular transformed proportion of soil variables of administrative villages in parentheses
[3]Odds ratio with 95% confidence interval in parentheses in ordinary logistic regressuion
[4]Crude odd ratios with 95% confidence interval in parentheses in multilevel logistic regressuion
[5]Adjusted Odd ratios for altitude, paddy field, dry crop land, forest with 95% confidence interval in parentheses in the multilevel logistic regression.

4.5 Discussion

4.5.1 Overview of the results

Malaria routine data, landscape and environmental characteristics were analysed in Yuanyang County, Yunnan, China. From 1987 to 1996, a total of 13,768 *P.vivax* and 1,874 *P. falciparum* and 11 mixed infection (*P.vivax* and *P. falciparum*) malaria cases were identified through routine malaria surveillance and reporting systems in the Yuanyang. The population and malaria cases in the towns and city were excluded from analysis to eliminate the possible bias of temporary migration that is more frequent from those areas. Malaria data and terrain maps, land use and soil maps were integrated into geographical information systems. The mean altitude and composition of different land uses and soils of all administrative villages were derived from the maps through integration of the maps with the administrative boundary map in GIS.

Logistic regression modelling was used to analyses risk factors of *P. vivax* and *P. falciparum* incidence in the administrative villages. The result of the analysis indicates that there were considerable spatial variations of both *P. vivax* and *P. falciparum* malaria in Yuanyang County, Yunnan. The spatial distributions of both *P. vivax* and *P. falciparum* infections in the county were highly associated with their landscape features and environment. The altitude of a village was negatively correlated with the level of both *P. vivax* and *P. falciparum* malaria. Paddy rice fields increased both the risk of *P. vivax* and of *P. falciparum* malaria. The amount of dry crop fields was negatively correlated with the level of *P. vivax* and *P. falciparum* incidence in the administrative villages. Forest tends to increase the relative risk of malaria in a village. The amount of bare rock in the village is also correlated with both *P. vivax* and *P. falciparum* malaria risk.

Multilevel logistic regression modelling was used for further analysis on removing possible *group effects* at township level in Yuanyang. The results of multilevel modelling showed more logical conclusions. The risk of *P.vivax* malaria was negatively correlated with the mean altitudes of the administrative villages. Paddy field was increasing the risks of malaria in the administrative village. More forest tends to increase the relative risks of *P. vivax* and *P. falciparum* infection, but the effect is not statistically significant. Similar trends were found in analysis of the *P. falciparum* malaria determinants. The associations of bush and bare rock and dry cropland with malaria were vanishing in multilevel analyses, which suggested that the apparent association of the three variables on the risk of malaria in ordinary logistic regression was due to the confounding "*group effect*" of administrative township. The multilevel logistic modelling also reveals the significant *group effects* (random parts) in the models, which could not be explained by the variables in the final logistic regression models for both *P. vivax* and *P. falciparum*.

4.5.2 Constraints of data and data analysis

One of the biggest problems of the analysis of routine data in epidemiological studies is the quality of data. The disease variations in the different places could be influenced by use of different diagnostic criteria. In the present study, strict diagnostic criteria were in use, i.e. only a patient with a febrile episode and malaria parasites detected by microscopic examination was considered to be a malaria case. The variation of disease in different places could be due to a different proportion of malaria cases reported. More malaria cases reported in one place might be due to more staff working in this place and/or more skilful technicians reading blood slides, or to the staff working more conscientiously. Multilevel modelling analysis was used to adjust for such variation, but it is not possible to correct for reporting variations among different village doctors within townships.

Here is an example to support this theory. A prospective cohort study was carried out in Feng Chun Ling Township, Yuanyang County in 1998. An active and passive malaria surveillance was carried out by our research team in a main transmission seasons to give accurate malaria incidence data. Detail of this study will be described in Chapter 5 of this thesis. We used the malaria incidence rates of 11 administrative villages in Feng Chun Ling from the survey of 1998 as a "golden standard". The mean malaria incidence rates of the 11 administrative villages through routine surveillance data from 1987-1996 were used to compare them with the "golden standard". Figure 4.18 shows the correlation between malaria incidence rates in the 11 administrative villages based on survey data in 1998 and the mean incidence from 1987 to 1996. Under the assumption that malaria is stable, underreporting is clearly shown as shown in the figures. Routine reporting data can only explain 16.7% (R^2=0.167) of total variation in the study area. Figure 4.19 shows the correlation between the indigenous malaria incidence rate in 1998 and the mean incidence from 1987 to 1996. The linear regression model can explain 45.7% of total variation. Therefore, although routine data can roughly reflect the malaria risk picture, they still serious biased in the Yuanyang County.

Figure 4.18. The correlation of malaria incidence rates based on survey data (1998) and routine surveillance data(1987-1996) in the 11 administrative village, Feng Chun Ling, Yuanyan, Yunnan

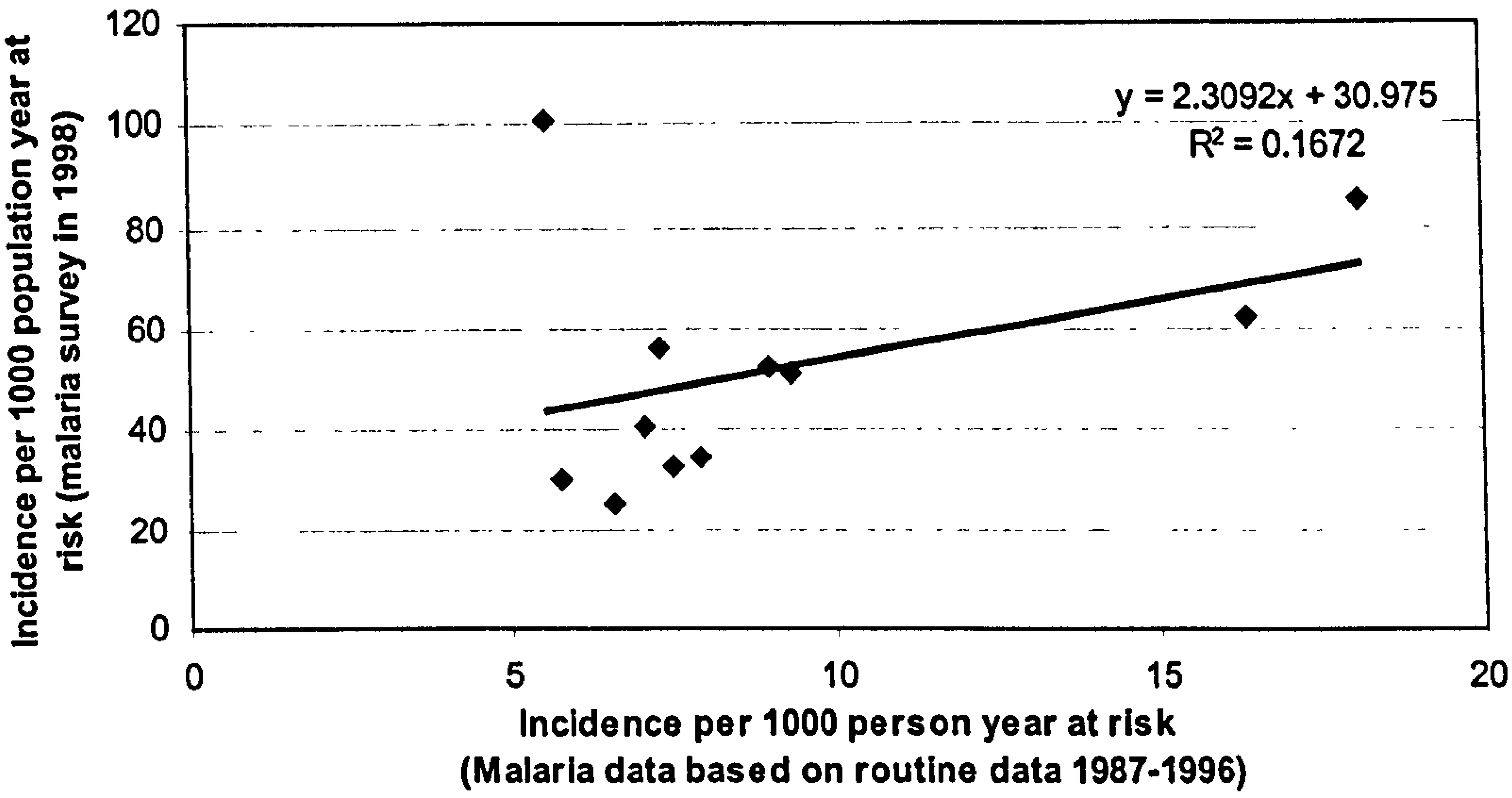
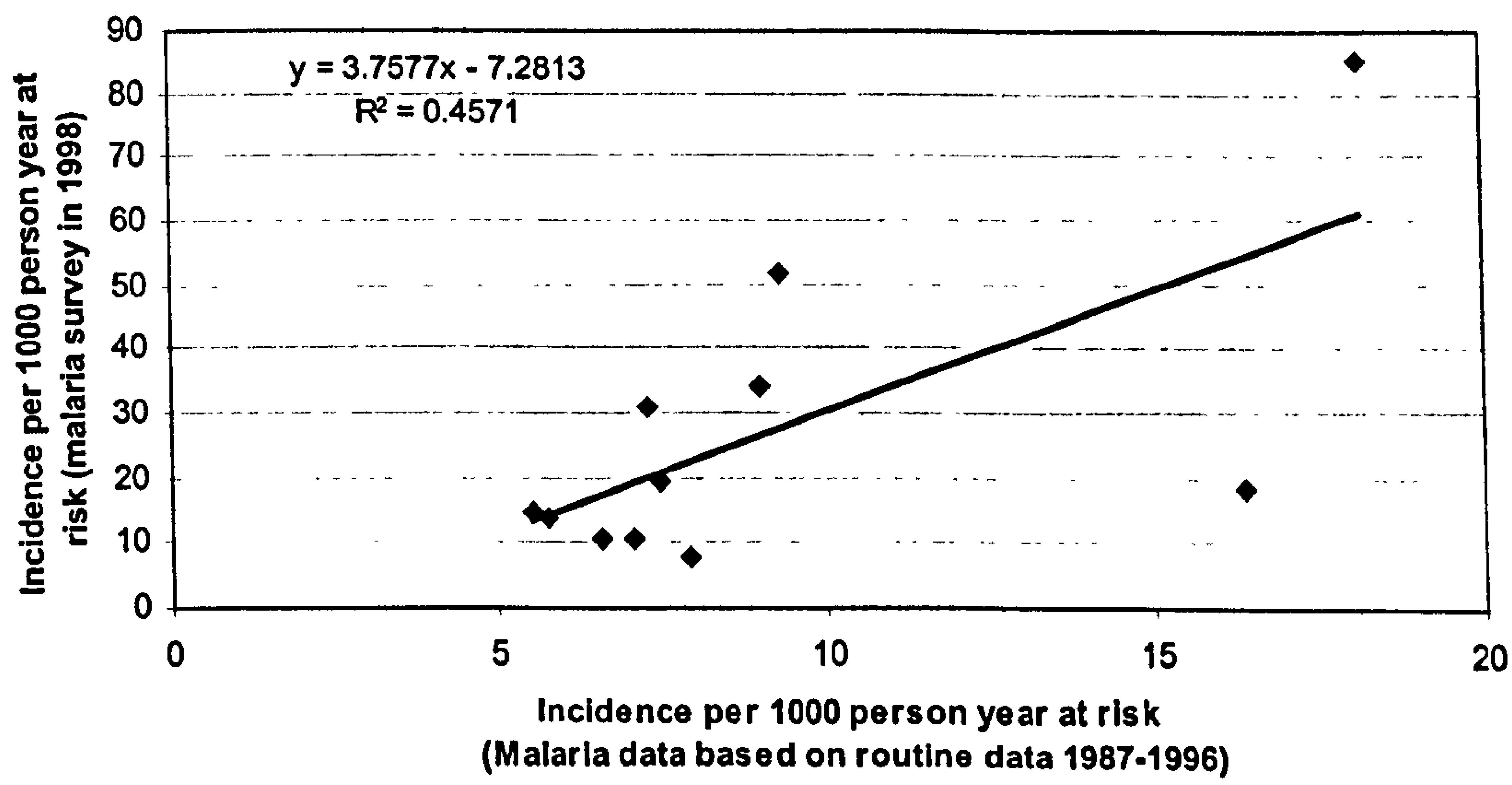


Figure 4.19. The correlation of indigenous malaria incidence rates based on survey data (1998) and routine surveillance data (1987-1996) in the 11 administrative village, Feng Chun Ling, Yuanyan, Yunnan



Malaria is unstable in most of Southeast Asia, particularly in Yunnan, China. Therefore, malaria tends toward outbreaks, annual variation of the disease in different places is obvious and it is very unstable in the study area. It is very difficult to assess the effect of spatial environmental variables on the risk of malaria in an unstable situation. We collected malaria routine data in one county with the population of 340,000 over a 10-year period to make the data stable. Geometric mean incidence rates were used in the hope of smoothing the malaria data for the analysis in the phase I study. In the multivariate regression analysis, a bi-variate outcome, high versus low incidence of *P. vivax* malaria in administrative villages, and present versus absent *P. falciparum* malaria infection, were used for the analysis to avoid too parsimonious a model such as the Poisson regression model as already discussed in 4.4.4

Routine data lacks geo-referenced accuracy. We tried to overcome this shortcoming by using relatively accurately geo-referenced units, administrative villages, as a unit of analysing the association between the risk of malaria and landscape and environmental features in the administrative village during the data analysis. The interpretation of the results must be within an assumption of homogenous spatial distribution of the population within administrative villages. It might be too arbitrary to assume the risk of malaria is homogenous in the administrative village, which is only ascribed to the landscape and environmental determinants within the boundary of the administrative village, particularly for those villages located near the boundary of an administrative village. Using a natural village or household as the unit of analysis would be more appropriate than using the administrative village as a unit for the analysis. The environmental determinants should be measured based on mosquito biology (flight range of mosquito) and malaria biology (transmission zone). Since malaria data are available through the routine reporting system at the level of administrative village, this analysis could help us to give a preliminary understanding of the importance of the landscape and environmental determinants of malaria.

The so-called "ecological fallacy" is one of the biggest problems in analysis of disease aetiology, as in present study. All the analyses are based on grouped malaria data and landscape and environmental data from the administrative villages instead of at the individual level. This fallacy cannot be eliminated in the phase I study. Interpretation of the results from an ecological study should be very careful. The results derived from an ecological study would not be valid for the smaller unit. In the present study, we cannot automatically assume the results are validated for individual, household, and even natural village levels. They are only validated for the administrative village.

Most confounding variables are not available for correcting the model for the present study. One of main confounding variables in this study is migration. To eliminate the migration bias, we only used malaria data from local residents and analysed the association of malaria incidence rates in local residents with various landscape and environmental variables. Non-permanent residents, the populations living in small towns and city were excluded from the analysis. Inter-village migration would be another problem in the analysis, that people who live at the top of mountain areas may have come down to the bottom of the mountain to plant tropical crops and fruits etc, but remain registered in their original residential areas. This misclassification of the origin of malaria infection would dilute the association of malaria incidence with altitude as discussed in section 4.4.5.

The last technical problem of the present phase I study is that the land-use map is out of date. The land use and land cover map was generated 1986, based on the survey from 1983 to 1985. In reality, the land-use pattern has changed greatly, particularly in lower altitude areas. A lot of land has been used for tropical fruits and crops etc. The land use type "paddy" does not change too much. More dry crop lands were generated in Yuanyang due to over population in the middle and top altitude areas in Yuanyang. Deforestation has occurred in the lower altitude area, and reforestation has occurred in the middle and high altitude areas in recent years. Consequently, it would affect the results of the modelling in the present study. We will use satellite images to update the land use map in further analysis for the main field study.

4.5.3 Data analysis

The number of the malaria cases and population sizes in each administrative village were collected in the phase I study. Ideally, continuous variables should be used to model the association between risk of malaria and environmental variables such as the multivariate linear regression model or the Poisson regression model, as they would be better able to capture more variation within data. However, more accurate models such as the Poisson regression model and multivariate linear regression model would give equal weight to each individual malaria case. Consequently, the model would be biased to the more accurately reporting villages and less weighting would be given to less complete malaria reporting villages. Furthermore, malaria is very unstable in Yuanyang County with outbreaks from time to time. Therefore the accurate models would be biased to high reporting villages and the villages that had had experienced outbreaks. To avoid this, logistic regression models were used in the present study. The villages were crudely stratified into two strata for *P. vivax* and two strata for *P. falciparum*. The stratification was based on the arguments that any village with malaria tends to report malaria although there is some variation in reporting. The crude stratification and modelling will ignore variation of malaria risk within strata, but it would give more valid results.

In summary, the phase I study indicated that malaria spatial distribution was highly varied in Yuanyang County. The spatial distribution of malaria in the county was highly correlated with 'altitude' and land use patterns such as "paddy" and "forest". There were a number of biases in this analysis, such as the quality of data, variations of malaria reporting in different areas, lack of information on potential variables for correcting the model, outdated land use map etc, and the study was pushing the limits of its "gold standard", the routinely recorded malaria cases. Therefore, a more sensitive and rigorous model needs a more intensive field study, and this will be dealt with in phase II (Chapter 5).

Chapter 5

Phase II main field study

5.1 Objectives

- To develop a predictive model of malaria spatial distribution in the study area based on:
- Creation of a geo-referenced data set of landscape, environmental factors and socio-economic variables related to malaria spatial distribution in the study area
- Quantification of malaria distribution related to landscape and environmental factors and socio-economic and human behaviour variables in the study area.

5.2 Specific hypotheses

Based on the findings of the phase I study, the following specific hypotheses were studied in the phase II, main field study

- ***People living at the lower altitude areas have a higher risk of malaria than those in higher altitude areas***
- ***People living in areas with more paddy rice fields around their households have a higher risk of infection with malaria.***
- ***People living in areas with more dry crops have a lower risk of malaria than those with less dry crop land.***
- ***People living near to or in forest areas tend to have a higher risk of* P. vivax *and* P. falciparum *infection; (the process of land resource development might increase the risk of malaria).***

5.3 Methodology

5.3.1 Study Design

The phase II study is a prospective follow-up study carried out in the field. Here I collected primary data rather than relying on routine data. A selected cohort was followed-up during the main malaria transmission season from May 1998 to December 1998. Malariometric variables, including malaria morbidity data, were collected during the follow up period. Meanwhile, socio-economic and behaviour data of the study population were also collected, and various landscape and environmental variables were derived from topographical maps and satellite imagery. The association between malariometric variables and different categories of landscape and environment were assessed to develop a predictive model.

5.3.2 Study area and population

The study area is located within Feng Chun Ling Township (Figure 5.1), in the Southeast of Yuanyang County, the Red River basin, Yunnan, China. The area was selected for the present study because

- The area is located within the county where the phase I study had been carried out, so that some data could be shared by both phase I study and phase II main field study
- The area is along the Red River with significant variations in altitude and is representative of much of the Red River basin area
- The area contains the mixtures of land use types representative of much of the Red River basin area.
- The area is co-endemic for *P. falciparum* and *P. vivax* and has a relatively high level of malaria incidence as determined in phase I (Figures 4.7 and 4.8).
- The area has a reasonably good and co-operative primary health care infrastructure and a better than average amount of health facilities.

The total population in the study area was around 32,000 based on the 1990 census. There are 104 villages in the area, which formed 13 administrative villages with an area of 18 x 18 km^2. Two administrative villages with 18 natural villages were excluded from the study due to the lack of co-operation by the village doctors, hence reliable data would not be available. Around 25,000 people within 86 villages in 11 administrative villages were included in the study. There are mixed land use types in the township (Figure 4.13), and a range of altitudes is significant within the study area (Figure 4.11). There is almost no flat land in the area. The vertical variation of climate is obvious, with a temperate zone above the altitude of 1,600 metres, the sub-tropical zone from 1,000 to 1,600, and the torrid zone below the altitude of 1,000 metres.

There are five ethnic minority groups and the Han Chinese group in the study area. The majority of the population farms mainly rice and dry crops. The people in the lowlands under the altitude of 800-900 metres also plant banana and other tropical fruits. The people in the high mountainous areas usually undertake agricultural work in the lowlands. Some of them work for mineral industries. Both lowland agriculture and mineral industries are located in the lower altitude areas with high malaria endemicity.

Figure 5.1. The geographical location of Feng Chun Ling, Yunnan, China

P. vivax and *P. falciparum* are co-endemic in the area. *An. minimus* and *An. sinensis* are the main malaria vectors. The clear and slow flowing terrace paddy fields provide the ideal breeding sites for *An. minimus* in the study area as in much of the Red River basin areas (Dong, 1993). Feng Chun Ling has been one of the most malarious areas in Yuanyang County since the middle of the 1980s (Figure 4.7 and Figure 4.8). The number of reported *P. vivax* and *P. falciparum* cases in each administrative village from 1987 to 1996 in the Feng Chun Ling township have ranked them among the most malarious villages in Yuanyang County. The geometric mean incidence rates of *P. vivax* and *P. falciparum* were 6.5 and 0.8 per thousand persons per year respectively in the township for the 10-year period, which is significantly higher than county-wide mean incidence rates. However, considerable under reporting of malaria cases is believed to have occurred in the area in recent years (Advisory Committee on Malaria, MoPH, 1994); it was not clear what is the exact level of malaria risk in the area. There is a township hospital located in the town of Feng Chun Ling. The disease control group within the hospital is responsible for malaria control and the malaria surveillance programme in the township. Each administrative village has its own village clinic, and some have more than one, mostly owned by the village doctors. Usually two village doctors work in a village clinic. In theory, the village doctors are responsible for treating clinical malaria cases in their villages, taking blood slides from febrile patients, and reporting the malaria cases to the township hospital. Their work sometimes overlaps among the villages.

5.3.3 Field operation and training

5.3.3.1 Field office

The phase II main field study was carried out in the main malaria transmission season during the rainy season from 10th May 1998 to 10th December 1998. A field office was set up for supervision of the implementation of the study headed by the author (***principal investigator***) in the Feng Chun Ling Township Hospital. He was assisted by a ***field co-ordinator***, who had had much experience in the management and technical aspects of a malaria project, and was recruited from Honghe (Red River) Prefecture Epidemic Prevention Station where he was also the head of the malaria control programme in the Prefecture. The office housed the laboratory for microscopic examination of blood smears, data management unit for the project, the whole database of fieldwork headed by a data manager and the computer unit for data entry. The office had sole and primary responsibility for the proper conduct of the study and held staff meetings weekly. The meetings were held to discuss progress and resolve any problems regarding fieldwork, surveillance reporting and quality of field workers' work.

5.3.3.2 Recruitment and training

Field Supervisors: Two field supervisors were recruited for the supervision of the fieldworkers to ensure they worked according to the research protocol of this study. Both of them were qualified medical doctors who had been responsible for malaria and other infectious disease control in Feng Chun Ling Township Hospital. The field supervisors were trained by the principal investigator (PI) with emphasis on how to conduct themselves, their job responsibilities and standardisation of fieldwork. Field supervisors participated in the weekly meeting held in the field office.

Field workers: All the field workers were recruited from local village doctors in the Feng Chun Ling Township. Most of the village doctors are responsible for malaria control and routine malaria surveillance and reporting, including taking blood slides from patients with a febrile episode during the main transmission season. An administrative village usually was provided with one or two field workers during the study period, depending on the size of the population and area covered by the administrative village. A total of 19 field workers were recruited. They received one week's intensive training to standardise the procedures of data collection and to make them aware of the regulation of fieldwork prior to starting. They were trained on (i) rules and regulations of field work, (ii) malaria and its diagnosis, and standardisation of the treatment of *P. vivax* and *P. falciparum*, (iii) how to obtain informed consent in the household, to carry out the census and to fill in questionnaires and forms properly, (iv) how to carry out weekly malaria surveillance and how to take blood slides, dry them, store and dispatch them, and how to behave themselves in the households and villages during their weekly surveillance. In addition, all doctors who were not recruited as field workers but practised medicine in the study area received two days training on how to make slides, to give standardised malaria treatment, and to fill in forms in case some febrile patients reported to their clinics. A bonus was paid to the doctors who sent slides to the field laboratory in the field office.

Laboratory technicians: Two technicians were recruited for the present study. One of them was from Yunnan Institute of Malaria Control (YIMC) and the other was from Yuanyang County Disease Control & Prevention Station. Both of them had more than twenty years' experience of reading blood slides and had well-known expertise in Yunnan. They were also provided training with emphasis on the way records and forms were to be filled. Their jobs were also checked by the PI and field co-ordinator.

Data manager and his assistants: A data manager was recruited for the management of all data in the study. All the questionnaires and forms from the study were sent to the data manager for checking completeness and accuracy. The data manager and his two assistants were responsible for coding all the questionnaires and forms of the study and they then sent them to the computer unit for data entry. The data manager and his two assistants were also trained by the PI on their responsibilities and how to code data and conduct their work. Their work was checked regularly by the field co-ordinator and PI.

Data entry clerks: A computer unit with two computers was set up by the PI and two data entry clerks were recruited for data entry. The two data entry clerks were computer illiterate before they were recruited. They were trained by the PI on how to use computers, and how to use EPI/INFO for the data entry. However data entry screens and check files were all generated by the PI. The data entry clerks were responsible for primary data checking but final cleaning of the data was carried out by the PI himself.

5.3.4 Census

A staggered household census was carried out in the 11 administrative villages in Feng Chun Ling Township at the beginning of May 1998. The aims of the census were to enumerate the population in the study cohort, and to record the baseline characteristics of each individual to be used to explain the malaria variation in the study area. At the stage of quantifying malaria environmental determinants, they could be used as the confounding variables to correct the model. The definition of a village is one sharing the same village leader; their inhabitants usually reside near together, spatially. A household was defined as one sharing the same kitchen. All members of the same household usually live in the same house structure. A household might have more than one house, but the houses usually are close together spatially. Only permanent residents who stay in the study area were included in the census. Those people working outside of their villages during the whole of the study period were be excluded from the study although they were "permanent" residents of the study village.

The first stage of the census was to obtain informed consent from local communities (villages), the heads of villages and then the heads of households and all individuals during the census. A specific meeting had been held for all leaders of villages in the township to obtain informed consent from them prior to the census. Informed consents were then obtained from the heads of all households and all individuals by field workers during the census. All individuals involved in the study understood fully the nature of the study, the reasons it was being undertaken, and possible benefit to them and their community as well as any possible harm they might suffer. Only those giving consent were recruited into the study cohort.

The face to face interview was used extensively in the census. Each individual was interviewed in detail by using standard questionnaires on demographic and socio-economic data. For children, mothers were used as respondents. On occasions, when the mother of children in the study was absent, the nearest care-taker or guardian of the child was used as a primary data source. The questionnaires included name, age, sex, ethnic group, education, use of mosquito nets and the names of the head of household and names of village. The quality of data collection was checked or double-checked randomly by field supervisors and/or the field co-ordinator and PI daily. If the questionnaire was not appropriately filled or had any omission or inconsistencies, the responsible investigator would be asked to correct it; if necessary, he or she would re-interview the subject.

5.3.5 Household geographical co-ordinates and their micro-environment

The geographical co-ordinates of all households and temporary huts were identified by use of a hand-held Global Positioning System (GPS) receiver (GPSIII, Garmin International) in the study area. Two single hand GPS receivers were used for this study. When a geographical location of a household was going to be detected, a GPS receiver was set up in the front of the target household or temporary hut. In most cases, the GPS unit was placed 5 metres in front of households. Five-minute-mean co-ordinate was recorded for each household with the aim of getting more accurate and stable position data. The co-ordinate system was recorded as latitude and longitude with decimal degrees and the map datum of the co-ordinate system was World Geodetic System 1984 (WGS 84), a common global reference system. Micro-environmental data of all households and temporary huts were also collected at the same time. The recording form included the name of village, name of head of household, latitude and longitude read from GPS of the household and the type of household. The data were re-checked to ensure the completeness of all sections and then coded by the data manager and his two-assistants, and submitted for data entry in computer unit.

5.3.6 Malaria surveillance

Malaria morbidity data during the study period were obtained through active and passive case detection systems. Thick and thin blood films were obtained from anyone with a history of a febrile episode (temperature ≥37.5° C) to check for malaria parasites during the study period. The recruited field workers carried out active malaria surveillance. A field worker was assigned a certain number of villages and households to carry out census and surveillance for all individuals in his/her target villages and households. They visited all their target households' residents weekly to collect blood films from anyone with current fever or history of fever in the last 7 days. Besides, they also recorded the number of days each person in the study cohort had been absent from the village during the previous week. The household to household survey was carried out by the same field worker until the end of the study. The village doctors and physicians in the township hospital formed the passive malaria detection system in the study area. Patients with febrile episodes presenting at village clinics, township hospitals and the field laboratory would have their blood slides taken to check for parasites. Two blood slides were taken from any febrile patient and one of them would be microscopically examined in the field laboratory, and the field co-ordinator or the principal investigator examined the other one. In addition to taking blood slides, the following data were collected: name, age, sex, detailed address of residence (names of village and head of household), methods of data collection (active or passive), any history of migration prior to present febrile episode during the past month etc. Symptomatic cases with *P. falciparum* positive blood films would be treated with artesunate, 600mg over 5 days, a new derivative of Qinqhaosu active against the erythrocytic stage of *P. falciparum* parasites (Klayman, 1985; Barradell & Fitton, 1995). Those with *P. vivax* positive blood films were treated with chloroquine 1.5g over 3 days and primaquine, 180 mg over 8 days, and the same dosages would be used for anti-relapse treatment for all *P. vivax* infection cases one month after. All the treatments were properly supervised to ensure compliance by patients. The field supervisors, field co-ordinator and PI would check field staff work through randomly visiting villages at different times of the study period to ensure that all staff had followed the procedures for data collection and sent antimalarials for anti-relapse treatment for malaria cases with *P. vivax* infection.

5.3.7. Map collection

Landscape and environmental data were obtained from various relevant authorities (some maps shared with the phase I study) during the study.

Administrative boundary map: The scale of the administrative boundary map is 1:100,000. The map has the detailed names and relative geographical location of all villages and the boundary of villages, described in section 4.2.3.2.

Topographical and terrain maps: The scales of maps are 1:50,000. The maps, based on the air photogrammetry survey in 1969 and produced by the Chinese Military Authority in 1978, were bought from Yunnan Provincial Bureau of Land Survey and Cartography. Two separate map sheets form the topographic and terrain map in the study area. The geographical co-ordinate system of the maps is Beijing 1954 Geographical Co-ordinate System with a latitude and longitude grid. The altitude contour system of the topographical maps is based on the Huanhai Altitude System. The contour interval of altitude of the maps is 10 metres.

Land use map. A land use map (1:150,000) of Yuanyang county was used to help to classify the satellite image, described in section 4.2.3.2.

5.3.8 Laboratory Methods

All blood slides were stained with Giemsa's stain; thin films were used to confirm species identification. All slides were coded, so that the technicians who examined the slides would be unaware of where the slides came from. The slides were examined under oil immersion for parasites (magnifying x1000 times) and 200 fields were read. The quality of blood slide examination was double-checked by the field co-ordinator and PI. All positive slides and around 10% of the negative blood slides chosen at random were re-examined by the field co-ordinator and PI during the study.

5.4 Data management and analysis

5.4.1 Criteria of diagnosis of malaria

Criteria for confirmed malaria diagnosis were patients with a history of fever or having an axillary temperature ≥ 37.5°C and malaria parasites positive on microscopic examination of thin and/or thick blood slides.

5.4.2 Data management

All census data and questionnaires were entered into EPI-INFO software during fieldwork. Before the data were entered into the computer, all data were first checked manually by the field supervisors, then the data manager and his two assistants, and then the field co-ordinator and PI. If the questionnaires and census data were wrongly filled in, the responsible field staff would be asked to verify and correct, and if possible, re-investigate as necessary. All individuals, households and villages were assigned a unique identification number. The data manager and his two assistants were responsible for coding all data. Data input fields had range checks and legal values to limit errors, and the data were double entered into two computers separately by two data entry clerks in the computer unit; any inconsistencies were detected and corrected. Data entry and cleaning were ongoing along with data collection in order to facilitate quality control and monitor the performance of field workers. Databases from the EPI-INFO were then converted into DBASE and incorporated into the Database Management System (DBMS) in the GIS software package ArcView or incorporated into the DBMS in ARC/INFO to link their co-ordinates with a common code.

The geographical co-ordinates of individual households were exported as a DBASE file and the Shapefile of a spatial coverage for all household created in ArcView. This Shapefile was then transferred to an ARC/INFO file for further spatial analysis. Multiple episodes from the same individual were considered to be separate cases in the database, provided they were distinct from each other. The criteria for the identification of distinct episodes were the same as in the pilot study described above (section 4.3.2). Mixed infections were counted both as a *P. vivax* and a *P. falciparum* infection.

5.4.3 Map digitising and cleaning

The 50 metre interval contours of topographical maps of the study area (1:50,000) were traced onto digitising sheets and digitised with the software package Cartalinx with a Summa graphics digitising table. All digitised sheets were registered to a master tic file, with a minimum of eight-tic positions for each sheet. The geographical co-ordinate system of the original topographical and terrain maps is Beijing 1954 Geographical Co-ordinate System. In order to have a consistent geographical co-ordinate system of the whole GIS data base in the study, the geographical co-ordinates of tic points of the Beijing 1954 co-ordinate systems of the terrain map were transformed to a latitude and longitude co-ordinate with map datum of WGS84 with the assistance of Dr Liu Bao from PLA Land Survey and Cartography College in Zheng Zhou, China. The decimal degrees of tic positions were input into the digitised maps with the co-ordinate system of the Latitude/Longitude. The accuracy of the digitising was assessed using the RMS. The RMS errors of 0.002 are acceptable when new data is digitised into existing coverage. If this RMS value was exceeded the tics were re-digitised. Digitised maps were imported into ARC/INFO for cleaning and editing. The attribute of each contour (altitude of the contour) was added into the coverage in ARC and coded numerically in ARCPLOT manually. After editing and cleaning or building individual coverage topology respectively in ARC and ARCEDIT, all coverages were then transformed to a UTM co-ordinate system in ARC/INFO.

The two separate terrain coverages were merged into a single terrain coverage of the study area in ARCEDIT to form the topographical features of the whole Feng Chun Ling Township. The terrain maps were cleaned as lines, and the TOPOGRID routine in ARC was used to transform the vector contour coverage of topographical maps into a continuous grid with cell size of 20 metres (which exactly matched the cell size of the SPOT image, see below). The mean altitude of each household was calculated per 5 x 5-pixel window within the GRID environment of ARC/INFO with the consideration of maximum position uncertainty of 30-100 metres of the single hand held GPS receiver, and then cross tabulating the new 5 x 5-pixel mean altitude grid with the Shapefile of geographical co-ordinates of all households in ArcView 3.0. The mean altitudes of the 5,007 households were read out from a cross-tabulation in ArcView, and then the mean altitudes of all households were entered into an Excel spread-sheet for further statistical analysis.

5.4.4 Satellite image classification

A satellite image was obtained from SPOT 4 (*Satellite Pour 1' Oberservation de la Terre, Mission 4*). The imagery was taken on 12th April 1999. The reason for choosing an image which was taken in later spring is because at this time different types of land use are much easier to distinguish in terms of their spectral differentiation in different bands of satellite image in the study area. Toward the end of spring, most of the paddy rice fields are full of water in the highlands or at the early stage of rice plantation in the lowlands; forests are green, and dry crop fields were bare or seeds had just been planted.

Satellite image classification is the process of sorting pixels (picture elements or cells) into a defined number of individual classes or categories, based on their data file values. If a pixel satisfies a certain set of criteria, then the pixel is assigned to the class that corresponds to the criteria. There are two ways to classify pixels into different categories: supervised and unsupervised classification. Supervised classification is the process where we select pixels that represent patterns of land use or land cover that can be recognised or identified. Then we "train" the computer system to identify pixels with similar characteristics (i.e. a set of signatures are derived from "training samples"). Thereafter, all pixels in the image will be sorted out into classes based on the signatures. The supervised classification is useful when we want to identify relatively few classes with distinct, homogeneous regions that represent each class (Lillsand & Kiefer, 1994; Mather, 1999). Unsupervised classification is a process in which the computer identifies statistical patterns inherent in the image, clusters all pixels that are statistically similar, and attaches them to individual classes. This approach is particularly useful in cases where there is little or no available observational or documentary evidence on the nature of the land cover types covered by the remote sensing image and where the analyst's proposed spectral classes are "mixed"(Mather, 1999).

An unsupervised classification algorithm was used to classify the SPOT 4 image in the present study, as the same class of land use and land cover might have mixed spectral classes at different altitudes and different stages of paddy and crop plantation. More than that, complicated topographical features might cause areas of the same land use to show completely different spectral features. It would be impossible to "train" a single or even a series of homogenous area(s) to represent single land use and land cover type in the study area. Unsupervised classification allows us to have more clusters than that of supervised classification, and therefore, to capture more spectral and spatial variation of land use and land cover in the study area.

Satellite image classification is much more complex in mountainous areas such as the present study area. The interaction of the angle and azimuth of the sun's ray with slope and aspect produces a *topographic effect* resulting in variable illumination. This impedes our ability to make use of images in mountainous areas. More accurately, the *topographic effect* can be defined as the difference in radiance value from an inclined surface compared to horizontal ones (Schneider & Robbin, 1998; Mather, 1999), the most serious being the shadowing effect. Band ratioing is one of most common ways to correct for the topographic effects (Mather, 1999), based on the principle that a certain component of reflectance in all bands is a result of the angular effect. By dividing one band by another, the uniform angular component is divided, leaving the variation that represents the difference in earth materials (Schneider & Robbin, 1998). We used near infrared (Band 3) and red (Band 2) bands of SPOT 4 image to created a new band of normalised difference vegetation index (NDVI) in the hope of eliminating the topographical effect, particularly the shadow effect. The calculation of NDVI is based on the formula NDVI= (Band 3 - Band 2)/(Band 3 + Band 2) (Lillsand & Killer, 1994). The index of NDVI reflects the relative amount of vegetation in pixels, thereafter, new NDVI band not only helps to correct the topographical effect but also helps differentiate land use types. Therefore, in the present classification, the NDVI band was stacked with Band 1, Band 2 and Band 3 together in the Stack Module of ERDAS Imagine 8.3.1(ERDAS, Atlanta, Georgia, USA) for further processing.

The SPOT image was a sub-set, which only covers the study area in Feng Chun Ling Township. In order to improve the spectral separability of land use types and obtain more accurate classification of the imagery, the sub-image was stratified into two strata, the highland and lowland based on a roughly 900-metre contour. The dominant land use types are forest and paddy in the highlands, and paddy, banana and other tropical plantation and bare land in the lowlands. The lowland stratum of the sub-image was further divided into two zones that used the Red River as a demarcation line for the two zones.

The three zones (one zone in the highlands and two lowland zones) were classified separately by satellite image process software ERDAS Imagine. The highland image was clustered into 45 clusters based on Band1, Band2, Band3 and NDVI using an ISODATA algorithm (Lillsand & Kiefer, 1994; Mather, 1999). Clusters were assigned into four classes of land use types, i.e. "forest", "paddy", "dry crop field", and "others" ("others" including, banana plantation and other tropical fruits, bare land etc.). The ways of assigning the land use and land cover were based on the reflectance of four bands of the images and old land use and land cover maps, interactively, spatially and spectrally. The two lowland images were clustered into 30 clusters. A similar methodology was employed to classify the two zones of the image in the lowlands into land use maps as those in the highlands. Then, the three zones of the two strata were merged to form a single land use map for the study area through spatial modelling in MODELER Module in ERDAS.

5.4.5 Geometric correction and resampling of satellite image

Satellite image data gathered by satellite sensors are the representation of the irregular surface of the earth. The image data themselves have no co-ordinate systems. Nevertheless, the images are distorted due to the curvature of the earth, relative movement of the satellite and the earth, and sensors being used. In order to be compared with other images or integrated with other GIS databases, the image must be changed to fit a map projection system or a reference system. Before rectifying the image data, we must determine the appropriate co-ordinate system for the database. In the present study, the Universal Transverse Mercator (UTM) co-ordinate system was used to make it compatible with other GIS database geographical projection systems in the study.

To rectify the satellite image, a set of ground control points (GCPs) has to be chosen. A GCP is a specific pixel in an image for which the output map co-ordinates are known. 18 GCPs, which could be identified in both the satellite image and on the ground, were used to rectify the present image. GCPs were taken in such a way as to make them evenly distributed in the study area. The GCPs include the intersection of two roads and two rivers, the intersection of a river and road, bridges and buildings. The co-ordinates of GCPs on the ground were either read from the topographical map or recorded using a GPS (10 minute-mean co-ordinate). The co-ordinates of the image data and the co-ordinates of GCPs obtained by GPS or reading from the topographical map were entered in the ERDAS Imagine through the keyboard. A second order polynomial equation was used to model the co-ordinates of images and GCP co-ordinates due to mountainous topographical features in the study area. The modelling results show that the random error (RMS) is 0.834. The RMS was less then one pixel (error of 16.7 metres), and was deemed acceptable in the present study. The detailed co-ordinates of GCPs and their correspondent co-ordinates in the image can be see in Appendix II.

The three methods most commonly used for image resampling are nearest neighbour, bilinear interpolation and cubic convolution. The nearest neighbour method, which simply takes the value of the pixel in the input values that are closest to the computed co-ordinates was used in this case. The method is fast and ensures that pixel values in the output image are "real" in that they are copied directly. Other methods, which change the pixel value, may create problem of integration, particularly in the case of thematic maps (Mather, 1999). Therefore, the nearest neighbour method was used in the present study.

5.4.6 Verification and accuracy of satellite image classification

The assessment of satellite image classification accuracy is a process of quantifying the errors due to incorrect labelling of pixels (Mather, 1999). Ideally, the assessment of satellite image classification should be on a pixel by pixel basis, i.e. the pixels of different classified elements are chosen randomly. Then, the sampled pixels are compared with those of the ground truth. But it would be very difficult to know exactly where the pixels are on the ground due to imperfect geometric correction and difficulty to access the pixel. In the present study, the RMS of the geometric correction was 0.834 pixels. We might think the GPS would help to identify the pixel location, but it would be very labour intensive even disregarding the maximum 30 to 100 metre error of the single hand GPS unit. Therefore, in the present study, we selected relatively large homogeneous land use types. The single hand GPS unit was used to identify the geographical co-ordinates in four or more corners of the land use training areas. The position of each corner was no less than 50 metres to the edge of testing area (ground truth). Ten-minute-mean co-ordinate for each point was recorded with latitude and longitude co-ordinates with map datum of WGS 84. Around 10 places for each category of land uses and land covers were surveyed. The larger the area in each category, the better the ground truth point. Due to the complicated terrain feature in the study area, the ground truth survey was chosen evenly for each land use type in the whole study area. The co-ordinates of ground truth of different land use types were imported into DBASE separately, and from there transferred into a Shapefile in ArcView. The Shapefiles were then transferred into ARC/INFO files. The points of co-ordinates were manually entered by mouse digitising to form polygons, which were then coded as correspondent land use types and rastered into a grid with a cell size of 20 metres, which made them exactly match the cell size of the SPOT image. The raster ground truth grids were overlain with geometrically corrected satellite images to help assess the accuracy of the satellite imagery classification in ERDAS.

The overall classification performance, as well as that of individual land use type, was evaluated. The overall classification accuracy was calculated by dividing the total number of correctly classified pixels by the total number of pixels of ground truth's of all the test sites. The individual classification performance was assessed by dividing the total number of correctly classified pixels of the specific land use type by the total number of pixels of ground truth of the specific land use type of the test sites.

5.4.7 Satellite image integration with other GIS database

After satisfying the accuracy of satellite image classification, the classified image was then added as a layer of GIS to form an updated land use map. The classified image was then transferred into a grid in the ARC/INFO. In order to obtain a household's environment, the proportion of each type of land use within a 1,000 metre buffer around the household was chosen to quantify the land use variables. The choice of a 1,000 metre buffer around a household is due to the consideration of the approximately 1,000 metre mean flight range of the major malaria vectors *An. minimus* and *An. sinensis* in the Red River basin area (Dong, 1993). In order to obtain each household's environment, separate grids for different land use types were created in the GRID. The FOCALSUM command in GRID was used to create a value for each cell, which represent the number of cells of a land use within 1000 metre buffer by using a 100 x 100 window (each cell in 20x20m), and the new grid coverage was then transferred into vector coverages in ARC/INFO. Household location coverage was overlaid with the land use vector coverage using the "INTERSECT" command in ARC, so that each individual household's micro-environmental data around the household within the distance of 1,000-metre buffer were generated. The INFO data were then exported to DBASE for further statistical analysis. The buffer proportions of all household were then normalised distribution by using angular transformation (arcsin $\sqrt{p}$).

5.4.8 Strategy of statistical analysis

The analysis of the phase II study used the integration power of GIS to generate landscape and environmental data for analysis. All malaria data, thematic maps, and satellite image data were integrated in the ARC/INFO environment with common co-ordinates. Overlay and buffering were used to link the environmental, ecological, and other demographic and socio-economic variables with both the incidences of *P. vivax* and *P. falciparum* malaria. GIS generated databases were input into statistical software for the statistical analysis. Tabulation of the risk of malaria with the different categories of socio-economic and behaviour variables and landscape and land use variables were used to see the general trend of the association of the risk of malaria and their determinants. Multivariate Poisson regression analysis was used to estimate the effect of the socio-economics, human behaviour, landscape and environment on malaria after controlling for other potential confounding variables. Multilevel modelling was used to correct for spatial autocorrelation of living in the same households and same villages to obtain better estimates. The regression and statistical analyses were carried out in STATA, *MLwiN* and S-Plus.

5.5Results

5.5.1 Basic characteristics of population.

In the 85 villages, a census revealed a total population of 24,280 living in 5,007 households in the study cohort. 732 people refused to participate in the study, mostly women, or men too old to response. The overall proportion of participants was 96.9%. The population of the villages in the study cohort ranged from 38 to 1,339 persons. On average, the village size was 285.7 persons. The family size in the study area also varied greatly, ranging from 1 to11. On average a household consisted of 4.9 persons. Those people who migrated into the study area after our census were not included in the study cohort. Children born during the study were also not included in the study cohort.

The basic characteristics of the population were surveyed during the census with the primary intention of enumerating the study population and identifying the socio-economic factors which might influence the transmission of malaria and its control and adjusting for any confounding effects on the landscape and environmental variables. The population age distribution in the study cohort is similar to that of the county-wide average during the 1990 population census (Table 5.1). The young population formed the majority in the study cohort. 60.41% were under 30 years old. 11.21% of the study population were pre-school children under 7 years old. One fifth of population was over 45 years old. This is the typical age structure in the less developed areas of China as well as other developing countries in the world. The age structure of the study cohort suggests that average life expectancy of the population in the study cohort was short.

There were more males than females (53.62% vs. 46.38%) in the study cohort. The overall male : female ratio is 1.16. The ratios, however, were different between age groups (Table 5.1). More boys than girls were in the young age group in the study cohort. But the number of population over 45 year old was roughly the same between the males and females. The sex ratio difference among the young age group might be due to family planning. On the other hand, the relatively fewer females in the middle age group might be due to a small proportion of women refusing to participate, among ethnic minority group women, and some young females tend to find a job in the towns and cities. They were not available for the study. In contrast, males in the high and middle mountain villages mostly got temporary jobs in local lower altitude areas, and were still available during the census. Therefore, fewer women than men in the young adult and middle age adult groups took part in the study.

Table 5.1. The age and sex distribution in the study cohort, Feng Chun Ling

Age groups	Male(%)	Female(%)	Total(%)	Sex ratio
1 - 6	1,566(12.03)	1,157(10.28)	2,723(11.21)	1.35
7 – 14	2,327(17.87)	1,987(17.65)	4,314(17.77)	1.17
15 – 30	4,141(31.80)	3,489 (30.99)	7,630(31.43)	1.19
31 – 44	2,627(20.18)	2,254 (20.02)	4,881(20.10)	1.16
>=45	2,359(18.12)	2,373 (21.07)	4,732(19.49)	0.99
All age groups	13,020(53.62)	11,260 (46.38)	24,280(100.00)	1.16

There are five ethnic groups in addition to the Han Chinese group in the study cohort, namely, Hani, Yi, Dai, Miao and Zhuang ethnic groups. The majority of the study population belongs to the Han Chinese (44.65%), Hani (39.48%) and Yi (10.62%) ethnic groups. The three other small groups, Dai (2.68%), Miao(2%) and Zhuang(0.57%), accounted for less than 6% of the total study cohort. The relative composition of the population of the ethnic groups in the study cohort was different from that of Yuanyang County as a whole as described in the phase I study. The majority of the population in the county as a whole was Hani (52.6%) and Yi (23.7%), and the Han Chinese only accounted for 12.4%, based on the data of the 1990 census. Apart from the Han Chinese group, all the minority ethnic groups tended to live together, spatially in the study area. It is very common that the entire population of a village belong to the same ethnic minority group in Feng Chun Ling. The spatial distribution of villages that were occupied by different ethnic groups, is quite obvious (Table 5.2). Hani and the Han Chinese are distributed in the whole range of altitude areas, but the majority of them reside in the middle and high altitude areas. Yi mostly live in the middle of the mountains. Dai, and Zhuang groups usually live in the lower altitude areas. The majority of Miao groups live in the lower altitude area although some of them live in the middle and high altitude areas.

More than half of the study population were illiterate (57.84%), without any formal schooling. Considerably more women than men were illiterate (73.01% vs. 44.65%) in the study cohort, which reflects the attitude of ethnic minority groups on their children' education. They pay more attention to their boys' education than their girls' education. Most of the illiterates were over 30 years old. Only around 6% of the population enjoyed more than secondary school education. 36.2% of the population had less than 5 years education at the elementary school level.

Table 5.2. The ethnic groups and Han Chinese and their altitude distribution in the study cohort, Feng Chun Ling

Altitude	Han	Hani	Yi	Dai	Miao	Zhuang	Total
<=800	126(1.16)	771(8.04)	38(1.47)	638(98.00)	72(14.85)	139(100)	1,784(7.35)
-1,200	1,629(15.03)	2,102(21.93)	1,996(77.39)	6(0.92)	244(50.31)	0	5,977(24.62)
-1,350	2,759(25.45)	1,926(20.09)	499(19.35)	7(1.08)	7(1.44)	0	5,198(21.41)
-1,500	4,526(41.75)	2,883(30.08)	40(1.55)	0(0.00)	61(12.58)	0	7,510(30.93)
>1,500	1,800(16.61)	1,904(19.86)	6(0.23)	0(0.00)	101(20.82)	0	3,811(15.70)
Total	10,840(100)	9,586(100)	2,579(100)	651(100)	485(100)	139(100)	24,280(100)

Using mosquito nets was not common in the study area. Less than one fifth of the study population (18.76%) owned mosquito nets. Nevertheless, the owners of mosquito nets were not homogeneously distributed in the whole area. Neither is the population though. The most mosquito net owners resided in the lower altitude area where the densities of mosquitoes were very high, and malaria transmission was also intensive. Other owners of mosquito nets were the people with a better socio-economic situation. The distribution of mosquito nets at different altitudes is shown in Table 5.3. Populations residing in the lower altitude area had significantly higher proportion of own mosquito nets ($X^2 = 3028, P < 0.0001$).

Table 5.3. The distribution of mosquito net ownership at different categories of altitude in the study cohort in Feng Chun Ling

Altitude	No net	Net	Total
<=800	769(43.11)	1,015(56.89)	1,784(100)
-1200	4,214(70.50)	1,763(29.50)	5,977(100)
-1350	4,330(83.30)	868(16.70)	5,198(100)
-1500	6,859(91.33)	651(8.67)	7,510(100)
>1500	3,552(93.20)	259(6.80)	3,811(100)
Total	19,724(81.23)	4,556(18.76)	24,280(100)

5.5.1 Environmental characteristics of households

The geographical co-ordinates of 5,007 households were identified by the two single-hand held GPS receivers in 85 villages from May to June 1998. The spatial distribution of the households is shown in Figure 5.2.

Most households were built with earth bricks and a straw roof (52.25%). 4.31% of the households were temporary huts with straw roofs and no wall. The rest of the households were built with cement bricks or regular board. The detailed quality of housing and the population distribution is shown in Table 5.4.

Table 5.4. Household characteristics and population distribution in the study area

Quality of household	No. household (%)	No. population(%)
Good [1]	2,175(43.44)	10,901(44.90)
Flat[2]	2,616(52.25)	12,540(51.65)
Temporary hut[3]	216(4.31)	839(3.46)
Total	5,007(100)	24,280(100)

[1]Built with cement bricks or regular board
[2]Built with earth brick and straw roof
[3]No wall usually with straw roof

5.5.3 Topographical and terrain distribution of households and their population

The terrain of the study area is highly variable (Figure 5.3) with the altitude ranging between 144 metres and over 2,500 metres above the sea level. The spatial distribution of the 5,007 households was categorised into 5 strata based on the mean altitude of the households. The detailed categorised results are shown in Table 5.5. As shown in the table, around one third of the study population resided in areas with an altitude below 1,200 metres, where malaria is highly endemic. More than half of the households were located in the area with an altitude above 1,300 metres, in which malaria transmission was less intensive, but most of the cases would be imported from the lower altitude area.

Table 5.5. Altitude distribution of households and population in Feng Chun Ling

Altitude	No. household	No. population
<=800	377(7.53)	1,781(7.34)
-1200	1,225(24.47)	5,977(24.62)
-1350	1,060(21.17)	5,198(21.41)
-1500	1,554(31.04)	7,511(30.93)
>1500	791(15.80)	3,813(15.70)
Total	5,007(100)	24,280(100)

Figure 5.2. Spatial distribution of households in the study area, Feng Chun Ling

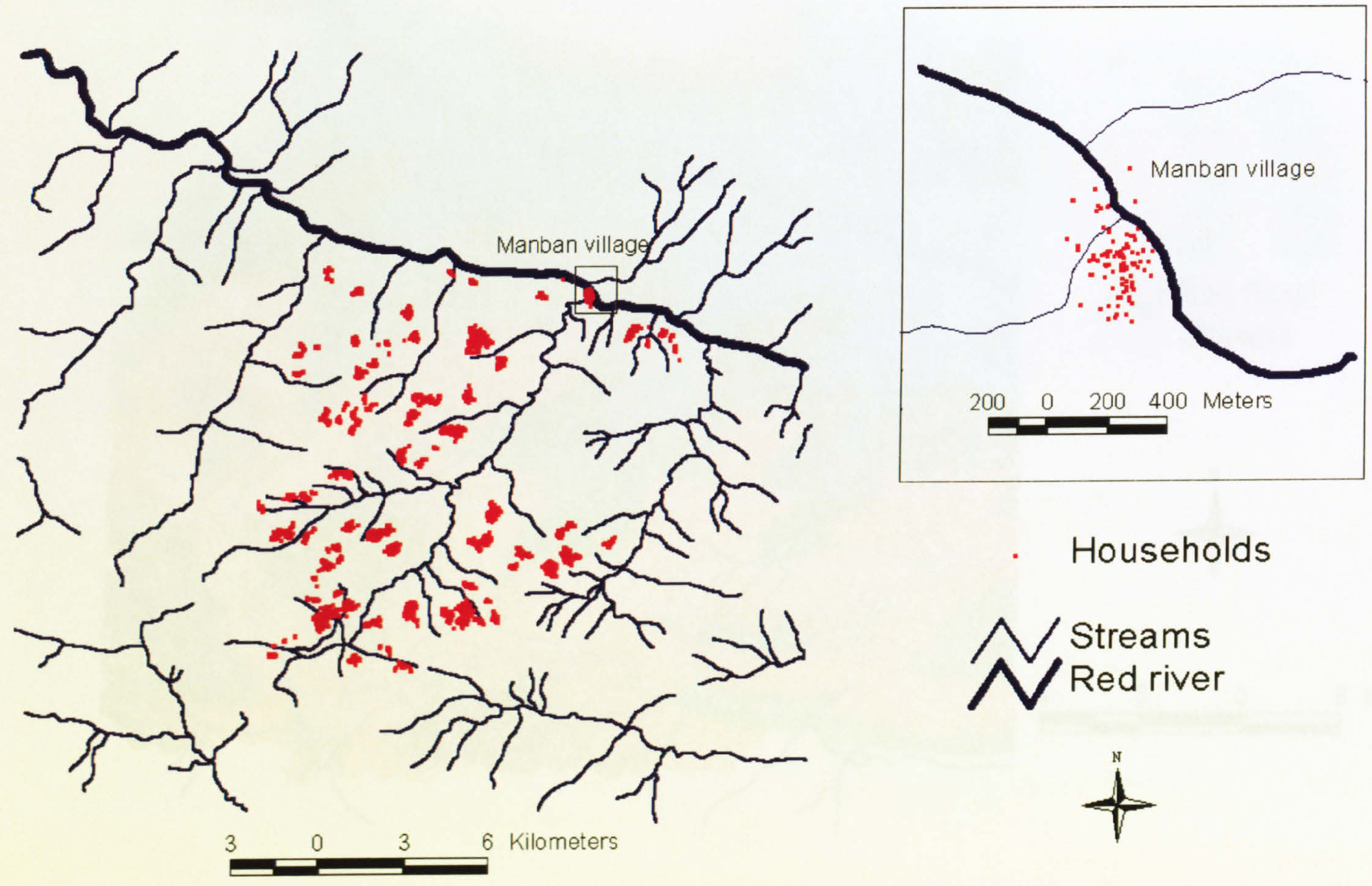
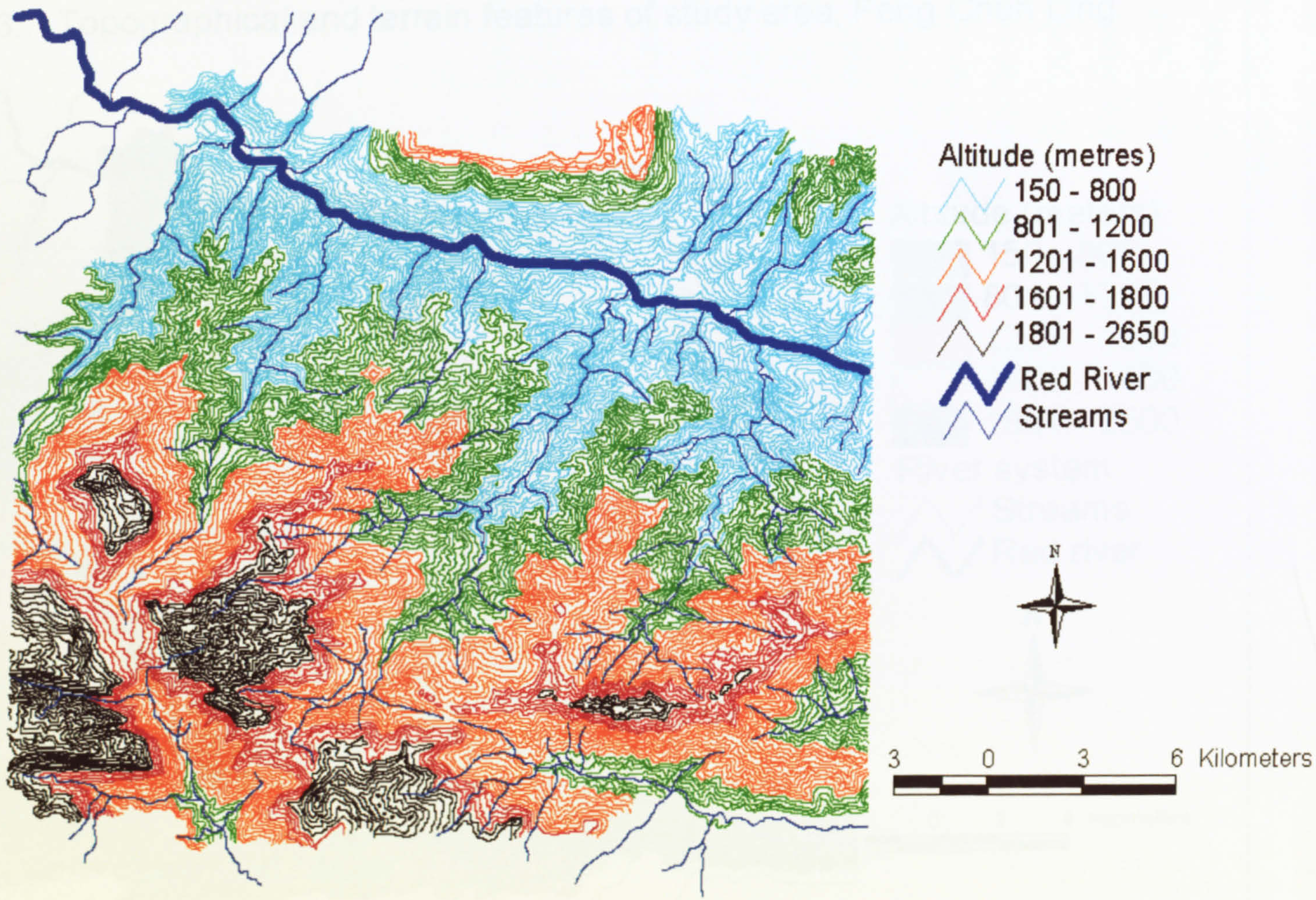


Figure 5.3. Topographical and terrain features of study area, Feng Chun Ling



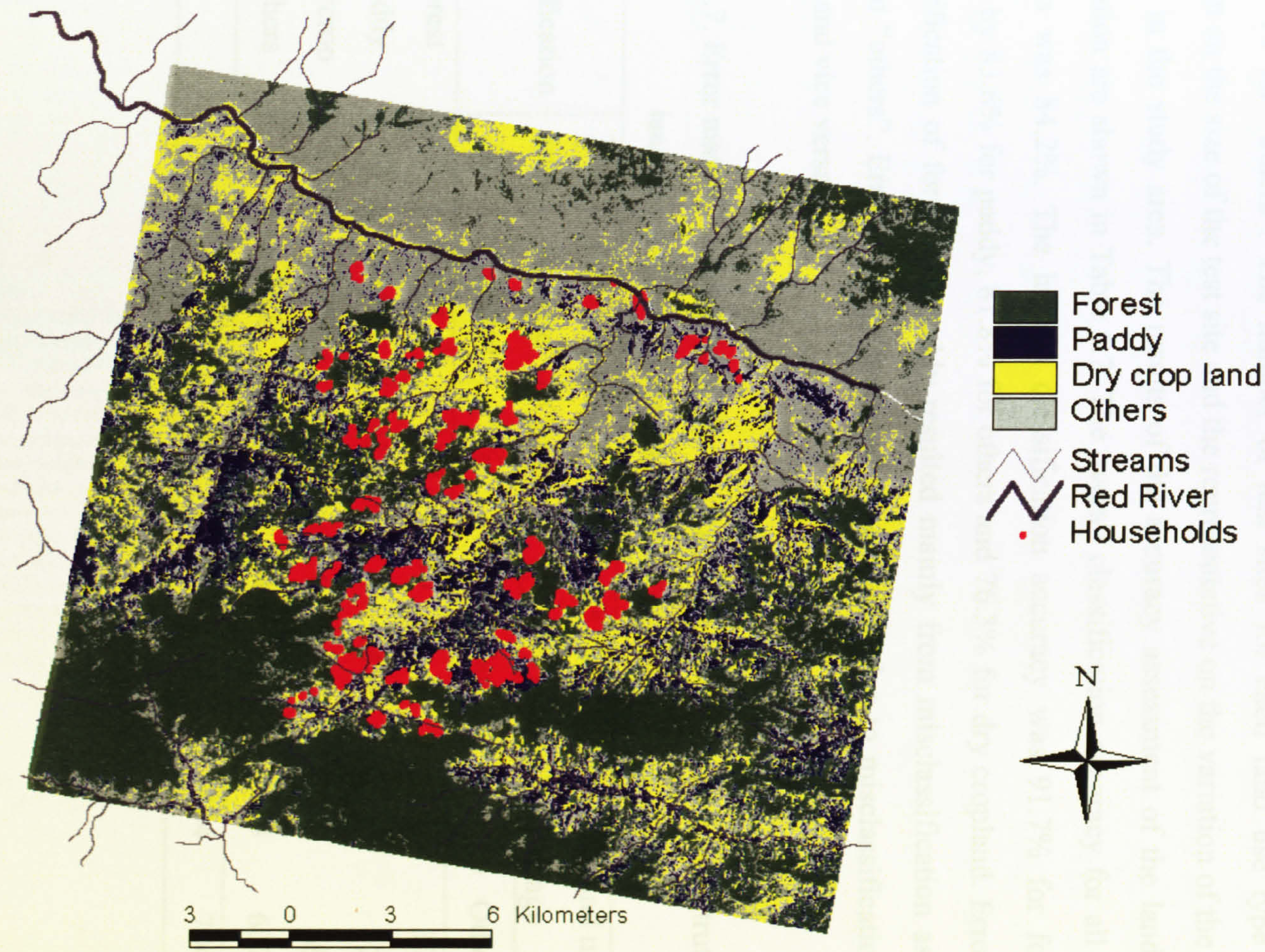
5.5.4. Classified satellite image and land use of households in the study cohort

The SPOT image was classified into four classes of land use types based on an unsupervised classification algorithm, namely, "forest", "paddy field", "dry cropland" and "others". The classification focused on the hypotheses of the main study. The cluster of "others" includes barren land, tropical fruit land, banana plantation, brush and bushwoods, roads and rock *etc.* Spatial patterns of the land use types and household locations are shown in Figure 5.4. The number of pixels and the proportion of land use types of the classified image were calculated in ERDAS. The result of satellite image classification showed that 31% of land in the study area was covered by forest, which is slightly higher than the county-wide average of land use map used in the phase I study (25.7%). Paddy field and dry crop field comprised 12.5% and 11.3% of the land of the study area, respectively, which is roughly similar to the county-wide average in the land use map described in the phase I study (Table 4.3). Table 5.6 shows the summary result of the sub-set of classified images that covered the study area.

Table 5.6 Summary of land use classification showing the total number of SPOT 4 image pixels, their equivalent area (km²) and the proportion of each of the land use elements over the study area covered by the sub-set classified image

Type of land use	No. pixels	Km²	Percentage
Forest	409,464	163.79	31.3%
Paddy field	162,987	65.19	12.5%
Dry crop field	147,378	58.95	11.3%
Others	586,971	234.79	44.9%

Figure 5.4. Land use pattern derived from SPOT image classification and household locations in the study area, Feng Chun ling



Overall 46 ground-truth test sites of land use types were surveyed, which consisted of 2,331 pixels corresponding to the sizes of SPOT pixel cells. The number of test sites used to determinate the classification accuracy varied by land use types, from 8 for forest to 14 for "others". The number of test sites for each land use type was dependent on the size of the test site and the representative on the variation of the land use type in the study area. The results of the accuracy assessment of the land use classification are shown in Table 5.7. The overall classification accuracy for all land use types was 84.2%. The highest classification accuracy was 91.7% for forest, followed by 83.6% for paddy, 81.8% for others and 76.3% for dry cropland. Errors in the classification of forest and paddy resulted mainly from misclassification as dry crops and "others". Errors for dry cropland were mainly due to misclassification as "others" and vice versa.

Table 5.7. Error matrix resulting from classifying SPOT 4 image and ground truth's land use test data in the study area, Feng Chun Ling

Classification	Test Sites (n)	Pixels in test site image(n)	Correct (%)	Pixels classified into the land use types in the SPOT image			
				Forest	Paddy	Drycrop	Others
Forest	8	671	91.7	615	7	12	37
Paddy	13	543	83.6	4	454	43	42
Drycrop	11	375	76.3	0	21	286	68
Others	14	742	81.8	31	11	93	607
Total	46	2,331	84.2	650	493	434	754

5.5.5 Association of altitude and land use variables in the study area

The relationships between altitude and the angular transformed land use variable were assessed. As shown in Figure 5.5, land use patterns around households are significantly correlated with altitude, particularly for the forest and "others". The relative amount of forest around households was significantly positively correlated with altitude ($r = 0.68$, $P<0.001$) and the land use variable "others" was significantly negatively correlated with altitude ($r = -0.73$, $P<0.001$) (Table 5.8). The results suggest that forest is mainly distributed in the higher altitude area and "other" mainly in the lower altitude area. As shown in Table 5.5.8, there was a significant negative correlation between paddy and forest ($r = -0.33$, $P<0.001$), and paddy and dry cropland ($r = -0.53$, $P<0.001$). A strong negative correlation between dry cropland and land use variable "others" ($r = -0.74$, $P<0.001$). The strong autocorrelation among landscape, and land use variables might make it very difficult to assess the effect of the variables on risk of malaria in univariate analyses.

Figure 5.5. The correlation of altitude w ith land use variables in study area

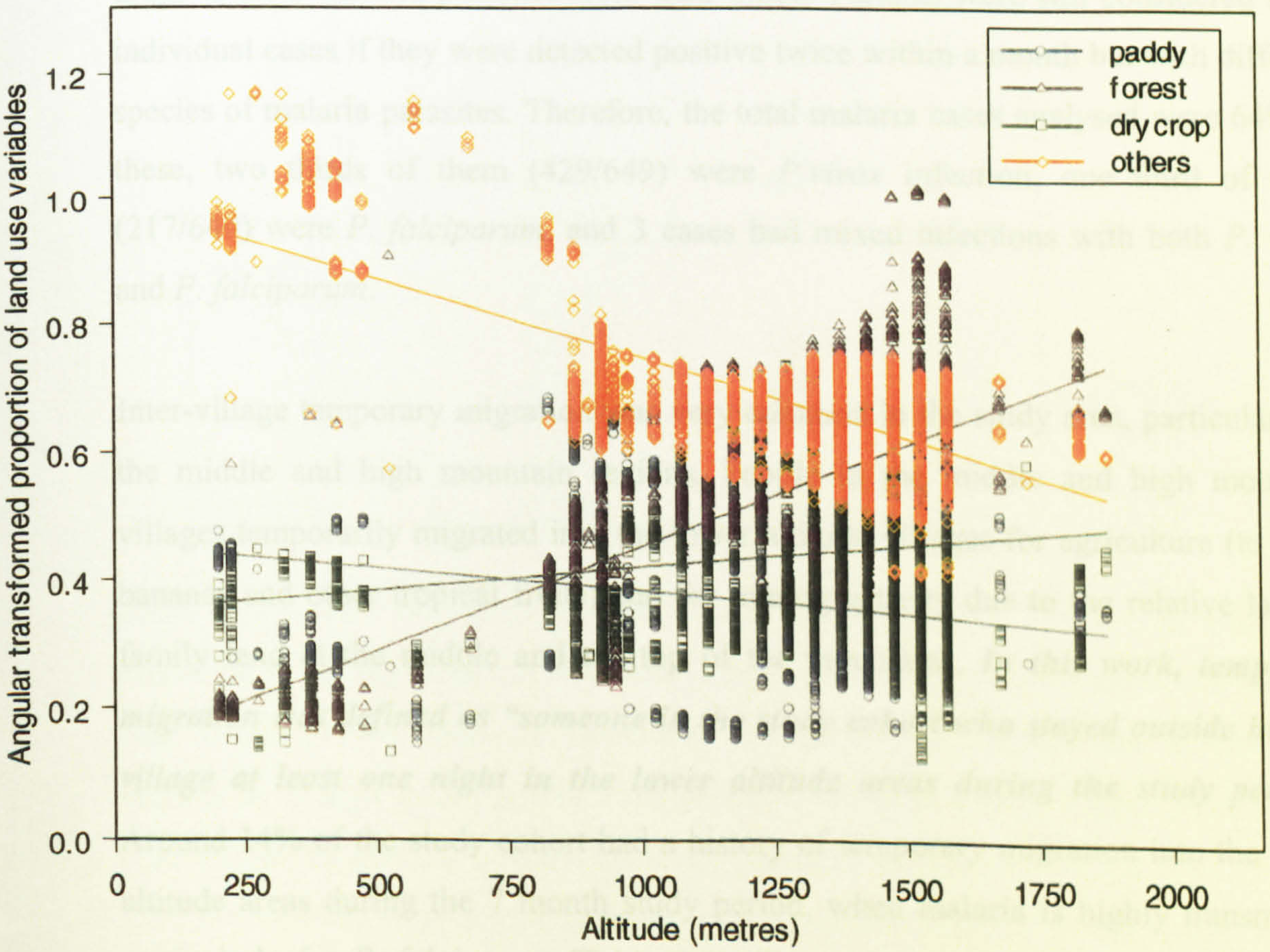


Table 5.8. Correlation matrix resulting from altitude, angular transformed proportion of paddy, forest, dry cropland and others in the study cohort

	Altitude	Paddy	Forest	Drycrop	Others
Altitude	1.00	-0.19	0.68	0.18	-0.73
Paddy	-0.19	1.00	-0.33	-0.53	0.12
Forest	0.68	-0.33	1.00	-0.19	-0.74
Dry crop	0.18	-0.53	-0.19	1.00	-0.26
Others	-0.73	0.12	-0.74	-0.26	1.00

5.5.6 Malaria data and population temporary migration

At the end of the study, a total of 5,360 slides had been collected from patients who had experienced febrile episodes during the seven-month study period. Of those 663 slides (12.37%) were identified as malaria parasite positive. Of the 663 malaria cases, 14 cases were excluded because they were identified as positive twice with the same species of parasites within one month. It was likely that the same patients saw different doctors for the same malaria episode or didn't take enough antimalarial drugs to clear malaria parasites from their blood. Patients were still considered to be individual cases if they were detected positive twice within a month but with different species of malaria parasites. Therefore, the total malaria cases analysed were 649. Of these, two thirds of them (429/649) were *P.vivax* infection, one third of them (217/649) were *P. falciparum*, and 3 cases had mixed infections with both *P. vivax* and *P. falciparum*.

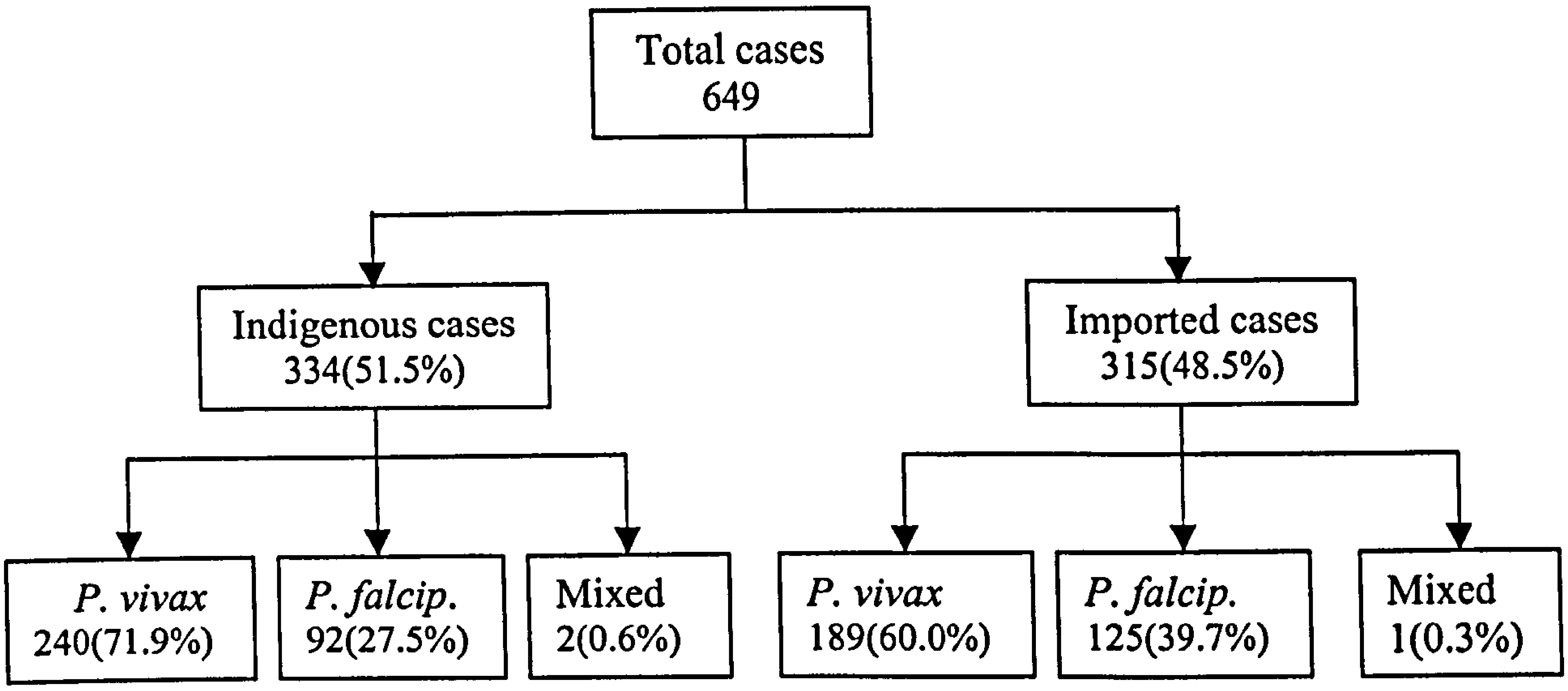
Inter-village temporary migration was very common in the study area, particularly in the middle and high mountain regions. People in the middle and high mountain villages temporarily migrated into the lower altitude villages for agriculture (to plant bananas and other tropical fruits) and for mining activity due to the relative lack of family land at the middle and the top of the mountains. ***In this work, temporary migration was defined as "someone in the study cohort who stayed outside his/her village at least one night in the lower altitude areas during the study period".*** Around 14% of the study cohort had a history of temporary migration into the lower altitude areas during the 7 month study period, when malaria is highly transmitted, particularly for *P. falciparum* (Table 5.9). The majority of the temporarily migrant populations were males (75%). The proportion of temporary migration among males was significantly higher than that among females *(X^2=701, $P<0.0001$)*. Around 80% of the temporarily migrating population were aged between 15 and 44 years old. And more than one in five of this age group in the study cohort had a history of temporary migration to the lowlands, for agriculture and other temporary jobs such as mining activity. A relatively small proportion of the elderly population and small children also migrated to the lowlands. The people most likely to migrate are from the area where the altitude was over 1,200 metres, and relatively fewer people in the low altitude areas were mobile (Table 5.9)

Table 5.9. The sex, age and altitude of household and the history of migration distribution of the study cohort

Variables	No. migration	Migration	Total	Proportion of migration(%)
Sex				
Male	10,500	2,520	13,020	19.4
Female	10,408	852	11,260	7.6
Sub-total	20,908	3,372	24,280	13.9
Age groups				
1 – 6	2,652	71	2,723	2.6
7 – 14	4,195	119	4,314	2.8
15 – 30	6,050	1,580	7,630	20.7
31 – 44	3,776	1,105	4,881	22.6
>=45	4,235	497	4,732	10.6
Sub-total	20,908	3,372	24,280	13.9
Altitudes				
<=800	1,783	1	1,784	0.1
-1200	5,579	398	5,977	6.7
-1350	4,344	854	5,198	16.4
-1500	6,140	1,370	7,510	18.2
>1500	3,062	749	3,811	19.7
Sub-total	20,908	3,372	24,280	13.9

Some cases were not infected within their native villages due to the temporary migration to the high malaria risk areas. ***The definition of a non-indigenous malaria case in the present study is "someone who in the past month had a history of staying in a high malaria risk area (the low altitude areas) 7 days prior to the current febrile episode".*** Based on this definition, around half of the total of cases (315 out of 649) were non-indigenous malaria cases (infected in the lower altitude areas) (Figure 5.6). The overall risk of indigenous malaria is 24.58 /1000 person years at risk based on present difinition. The risk for *P.vivax* was 17.81/1000 person years at risk and for *P. falciparum* was 6.92 /1000 person years at risk. Nevertheless, the present definition might misclassify some imported malaria cases with an incubation period longer than one month into indigenous malaria cases.

Figure 5.5. Malaria case classification based on the assumed source of infection



The proportion of *P.falciparum* infection of the imported case group was much higher than that of the indigenous group. The difference between the two groups was statistically significant ($X^2=10.59$, $P=0.001$). This reflects that the imported cases get their infection in the lower altitude areas where *P. falciparum* was the dominant infection among the inhabitants.

Table 5.10 shows the distribution of indigenous malaria and imported malaria cases based on the present indigenous malaria case definition in the study cohort (mixed infection counted as both *P. vivax* and *P. falciparum* infection). The majority of the imported cases were from the age group 14 to 44 in the highlands, who also migrated the most. Only a small proportion of malaria cases was from children and the older population. Over 90% of the imported cases were from the areas with altitude above 1,200 metres, where also the most active migration occurred due to scarcity of land in the areas.

A further crude analysis was carried out to compare the risk of malaria of temporarily migrating population with non-migrant population in the study cohort regardless of whether the malaria cases were imported or indigenous malaria cases as shown in Table 5.11. The overall risk of people with and without a temporary migration history were 118.6 and 12.1 per 1000 persons, respectively, during the 7 month study period. The relative risk of the migrant population against non-migrant population was 9.8, suggesting the migrant population had around 10-fold higher the risk of malaria than those of the non-migrant population. But there were large variations in the relative risks by sex, age as well as altitude as shown in Table 5.11. Among migrants the relative risks of children were much higher than of adults, and people in the higher mountain areas also had higher relative risk as shown in the Table. These imply that young age groups and the population in higher altitude areas were more vulnerable due to lacking immunity to malaria. The reason that migrant females had a higher relative risk is not very clear, it might be also because the females were less mobile and thus less immune to malaria. Although this crude analysis has a number of methodological problems such as comparability of the two groups and that the relative small number of people in some categories made the ratios unstable. But this crude analysis demonstrates that migration was a very important malaria indicator in the study area.

Table 5.10. Age and altitude distribution of indigenous malaria and imported malaria cases in the study cohort in Feng Chun Ling*

	Indigenous cases			Imported cases		
Variables	Male	Female	Total	Male	Female	Total
Ages (Year)						
1 – 6	18	14	32	9	11	20
7 – 14	33	38	71	27	10	37
15 – 30	81	42	123	120	47	167
31 – 44	27	42	69	39	26	65
>=45	24	17	41	19	8	27
Total	183	153	336	214	102	316
Altitude (m)						
<=800	48	32	80	0	0	0
-1200	56	35	91	17	2	19
-1350	30	28	58	58	31	89
-1500	34	43	77	90	43	133
>1500	15	15	30	48	27	74
Total	183	153	336	213	103	316

*Mixed infection counted as both *P. vivax* and *P. falciparum* infection

Table 5.11. The relative risk of malaria among migrated and non-migrated population groups in the study cohort, Feng Chun Ling.

Variables	No migration			Migration			Ratios*
	No. pop.	No. cases	Incidence/1000	No. pop.	No. cases	Incidence/1000	
Sex							
Male	10,500	137	13.1	2,520	260	103.9	7.9
Female	10,408	115	11.1	852	140	164.3	14.8
Total	20,908	252	12.1	3,372	400	118.6	9.8
Ages							
1 – 6	2,652	31	11.7	71	21	295.8	25.3
7 – 14	4,195	60	14.3	119	48	403.3	28.2
15 – 30	6,050	81	13.4	1,580	209	132.2	9.9
31 – 44	3,776	44	11.7	1,105	90	81.4	7.0
>=45	4,235	36	8.5	497	32	64.4	7.6
Total	20,908	252	12.1	3,372	400	118.6	9.8
Altitude							
<=800	1,783	80	44.9	1	0**	0.0	0.0
-1200	5,579	79	14.2	398	31	77.9	5.5
-1350	4,344	30	6.9	854	117	137.0	19.9
-1500	6,140	47	7.7	1,370	163	119.0	15.5
>1500	3,062	16	5.2	749	89	118.8	22.8
Total	20,908	252	12.1	3,372	400	118.6	9.8

*Ratio is the incidence rate of migrant population over that of no migrant population
** This group is effectively non-existent because of the definition of migration

5.5.7. Malaria risk for different demographic and behavioural variables

The distribution of indigenous malaria with different demographic variables is shown in Tables 5.12 and 5.13. The risk of malaria among males and females was roughly similar although females were slightly at less risk than males. It is probable that males had more exposure to mosquitoes due to temporary migration from one place to another for work. The age distribution of malaria showed the classic age distribution of malaria in Asia. Malaria basically infected all age groups. The school children, young adults and adults had a relatively higher risk of malaria infection. Pre-school children and older people generally had less risk of malaria. That might be due to the fact that young adults and the adult group tended to be more active at night. Hence, they are more exposed to mosquitoes. Young adults and adults were also active migrants to the low altitude areas. They were very much more likely to get the infection from outside of their villages.

Malaria affected all ethnic minority groups, but the Dai, Miao and Zhuang ethnic groups were at much higher risk than others. That might be because the three groups usually lived in the lower altitude areas in which the higher temperatures favoured mosquito breeding and malaria transmission as shown in Table 5.2.

The relationship between the number of years of education and the risk of malaria is also shown in Tables 5.12 and 5.13. It seems that more education is associated with a higher risk of malaria. That is probably because the educated populations were more sociable and liable to go out during the night. Another interpretation is that more educated people tended to see their doctors often, but this is unlikely due to the good surveillance was set up in the study area. It might be due to chance.

It has long been known that that mosquito nets tend to be protective from mosquito biting and malaria infection, but present data suggest that people using mosquito nets have a significantly higher risk of malaria (Tables 5.12 and 5.13). That may be because people who live at lower altitudes with the highest density of mosquitoes and risk of malaria were more likely to own mosquito nets for personal protection from mosquito biting as shown in Table 5.3. The univariate analysis could not identify the association between mosquito nets and risk of malaria since other landscape and environmental variables were so strong in determining the malaria spatial distribution in the study area, particularly altitude as shown in Tables 5.14 and 5.15.

People with a history of travelling outside their village had much a higher risk than those people who never moved outside their villages, both for *P. vivax* and *P. falciparum* infection (Tables 5.12 and 5.13). Most travellers went to higher risk malaria areas for agriculture and mining work. The place was very hot, and they mostly lived in temporary huts, which are easily accessible to mosquitoes. Consequently they became infected with malaria parasites. This supports our theory that our original definition of indigenous malaria cases might misclassify some imported malaria cases with an incubation period longer than one month into indigenous malaria cases as described in section 5.5.6 of this Chapter.

Table 5.12. *P.vivax* indigenous malaria distribution based on demographic, socio-economic and behaviour variables in the study cohort of Feng Chun Ling

Variables	No. pop.	PYAR[1]	No. vivax	Incidence per 1000 person year at risk[5]
Sex				
Male	13,020	7,130.9	128	17.95(15.10, 21.35)
Female	11,260	6,459.7	114	17.65(14.69, 21.20)
Ages				
1 - 6	2,723	1,584.4	21	13.25(8.64, 20.33)
7 – 14	4,314	2,520.0	58	23.02(17.79, 29.77)
15 – 30	7,630	4,158.0	84	20.20(16.31, 25.02)
31 – 44	4,881	2,645.5	49	18.52(14.00, 24.51)
>= 45	4,732	2,682.7	30	11.18(7.82, 15.99)
Ethnic Groups				
Han	10,840	6,042.0	78	12.91(10.34, 16.12)
Hani	9,586	5,304.0	124	23.38(19.61, 27.88)
Yi	2,579	1,500.6	16	10.66(6.53, 17.40)
Dai	651	382.4	15	39.23(23.65, 65.07)
Miao	485	279.8	7	25.02(11.93, 52.47)
Zhuang	139	81.8	2	24.45(6.11, 97.75)
Education				
Illiterate	14,044	7,990.1	123	15.39(12.90, 18.37)
Elementary	8,790	4,828.2	100	20.71(17.03, 25.20)
Secondary	1,446	772.3	19	24.60(15.69, 38.57)
Net				
No	19,724	1,1017.4	187	16.97(14.71, 19.59)
Yes	4,556	2,573.2	55	21.37(16.41, 27.84)
Quality of house				
Temporary hut[2]	839	460.7	12	26.05(14.79, 45.86)
Flat[3]	12,540	7,030.6	112	15.93(13.24, 19.17)
Good[4]	10,901	6,099.3	118	19.35(16.15, 23.17)
Migration				
No	20,908	12,315.7	187	15.18(13.16, 17.52)
Yes	3,372	1,275.0	55	43.14(33.12, 56.19)

[1] Person years at risk
[2] No wall usually with straw roof
[3] Built with earth brick and straw roof
[4] Built with cement bricks or regular board
[5] 95% confidence interval in parentheses

Table 5.13. *P.falciparum* indigenous malaria distribution based on demographic, socio-economic, and behaviour variables in the study cohort of Feng Chun Ling

Variables	No. pop.	PYAR[1]	No. falcip.	Incidence per 1000 person year at risk[5]
Sex				
Male	13,020	7,130.9	55	7.71 (5.92, 10.05)
Female	11,260	6,459.7	39	6.04 (4.41, 8.26)
Ages				
1 - 6	2,723	1,584.4	11	6.94(3.85, 12.54)
7 – 14	4,314	2,520.0	13	5.16(2.99, 8.88)
15 – 30	7,630	4,158.0	39	9.38(6.85, 12.84)
31 – 44	4,881	2,645.5	20	7.56(4.88, 11.72)
>= 45	4,732	2,682.7	11	4.10(2.27, 7.40)
Ethnic groups				
Han	10,840	6,042.0	36	6.79(4.90, 9.41)
Hani	9,586	5,304.0	24	3.97(2.66, 5.93)
Yi	2,579	1,500.6	6	4.00(1.80, 8.90)
Dai	651	382.4	22	57.53(37.88, 87.37)
Miao	485	279.8	3	36.67(11.83,113.69)
Zhuang	139	81.8	3	10.72(3.46, 33.24)
Education				
Illiterate	14,044	7,990.1	50	6.26(4.74, 8.25)
Elementary	8,790	4,828.2	36	7.46(5.38, 10.34)
Secondary	1,446	772.3	8	10.36(5.18, 20.71)
Net				
No	19,724	11,017.4	60	5.45(4.23, 7.01)
Yes	4,556	2,573.2	34	13.21(9.44, 18.49)
Quality of house				
Temporary hut[2]	839	460.7	37	6.07(4.40, 8.37)
Flat[3]	12,540	7,030.6	52	7.40(5.64, 9.71)
Good[4]	10,901	6,099.3	5	10.85(4.52, 26.07)
Migration				
No	20,908	12,315.7	65	5.28(4.14, 6.73)
Yes	3,372	1,275.0	29	22.75(15.81, 32.73)

[1] Person years at risk
[2]No wall usually with straw roof
[3]Built with earth brick and straw roof
[4]Built with cement bricks or regular board
[5]95% confidence interval in parentheses

The risks of *P. vivax* and *P. falciparum* indigenous malaria for different categories of altitude and land use variables are shown in Tables 5.14 and 5.15. There were large variations in the risk of malaria by altitude. There were clear negative associations of *P. vivax* and *P. falciparum* malaria with altitude in the study cohort. The higher the altitude, the lower the risk of malaria. The angular transformed proportion of land use variables: "forest", "paddy", "dry crop land" and "other" were categorised into quintiles in this analysis so that the same number of subject in each category would be achieved. This is quite a sensible way of choosing the group intervals provided the actual intervals are reported, although such intervals will vary from study to study, thus making it harder to compare findings (Clayton & Hill, 1993). There was a negative trend of the risk of *P. vivax* and *P. falciparum* and the relative amount of forest around household in the study area. However, the amount of forest is positively correlated with altitude ($r = 0.68$ $P<0.001$) and negatively correlated with paddy ($r = -0.33$ $P<0.001$) in the study area as shown in Table 5.8. Therefore, the relationship between forest and risk of malaria cannot be identified by the crude analysis. The risks of *P. vivax* and *P. falciparum* for the different categories of paddy showed high variations. There were general trends that more paddy was associated with a higher risk of both *P. vivax* and *P. falciparum.* There was a negative trend of the risk of malaria and the relative amount of dry cropland around households. This might also be due to the confounding effect of other variables, particularly paddy ($r = -0.53$ $P<0.001$). The risk of *P. vivax* and *P. falciparum* with the different categories of land use variable "others" showed large variation, but no general trend could be identified.

Table 5.14. *P.vivax* indigenous malaria distribution based on altitude and land use variables in the study cohort of Feng Chun Ling

Variables[1]	No. pop.	PYAR[2]	No. vivax	Incidence per 1000 person years at risk[3]
Altitude				
<=800	1,784	1,049.1	43	40.99 (30.40, 55.27)
~1200	5,977	3,439.2	70	20.35(16.10, 25.73)
~1350	5,198	2,881.5	52	18.05(13.75, 23.68)
~1500	7,510	4,133.0	58	14.03(10.85, 18.15)
>1500	3,811	2,087.8	19	9.10(5.81, 14.27)
Forest				
1st quintile(0.166-0.404)	4,856	2,787.0	73	26.19(20.82, 32.95)
2nd quintile(0.404-0.507)	4,856	2,701.6	44	16.29(12.12, 21.89)
3rd quintile(0.507-0.567)	4,856	2,674.2	39	14.58(10.66, 19.96)
4th quintile(0.567-0.671)	4,856	2,676.1	53	19.81(15.13, 25.92)
5th quintile(0.671-1.019)	4,856	2,751.8	33	11.99(8.53, 16.87)
Paddy				
1st quintile(0.148-0.279)	4,856	2,759.4	35	12.68(9.11, 17.67)
2nd quintile(0.279-0.325)	4,856	2,649.9	39	14.72(10.75, 20.14)
3rd quintile(0.325-0.381)	4,856	2,696.6	32	11.87(8.39, 16.78)
4th quintile(0.318-0.455)	4,856	2,715.4	59	21.73(16.84, 28.04)
5th quintile(0.455-0.732)	4,856	2,769.4	77	27.80(22.24, 34.76)
Dry crop land				
1st quintile(0.128-0.332)	4,856	2,749.9	70	25.46(20.14 32.18)
2nd quintile(0.332-0.377)	4,856	2,744.4	55	20.04(15.39 26.10)
3rd quintile(0.337-0.439)	4,856	2,755.5	55	19.96(15.32 25.99)
4th quintile(0.439-0.566)	4,856	2,729.6	31	11.36(7.99 16.15)
5th quintile(0.566-0.735)	4,856	2,611.1	31	11.87(8.35 16.88)
Others				
1st quintile(0.412-0.623)	4,856	2,677.1	36	13.45(9.70, 18.64)
2nd quintile(0.623-0.658)	4,856	2,710.1	57	21.03(16.22, 27.27)
3rd quintile(0.658-0.688)	4,856	2,698.2	50	18.53(14.05, 24.45)
4th quintile(0.688-0.718)	4,856	2,704.8	40	14.79(10.85, 20.16)
5th quintile(0.718-1.171)	4,856	2,800.4	59	21.07(16.32, 27.19)

[1]Angular transformed proportion of land use variables in parentheses
[2]Person years at risk
[3]95% confidence interval in parentheses

Table 5.15 *P.falciparum* indigenous malaria distribution based on altitude and land use variables in the study cohort of Feng Chun Ling

Variables[1]	No. pop.	PYAR[2]	No. falcip.	Incidence per 1000 per years at risk[3]
Altitude				
<=800	1,784	1,049.1	37	35.27(25.55, 48.68)
~1200	5,977	3,439.2	21	6.11(3.98, 9.37)
~1350	5,198	2,881.5	6	2.08(0.94, 4.64)
~1500	7,510	4,133.0	19	4.60(2.93, 7.21)
>1500	3,811	2,087.8	11	5.27(2.92, 9.51)
Forest				
1st quintile(0.166-0.404)	4,856	2,787.0	44	15.79(11.75, 21.22)
2nd quintile(0.404-0.507)	4,856	2,701.6	15	5.55(3.37, 9.21)
3rd quintile(0.507-0.567)	4,856	2,674.2	14	5.24(3.10, 8.84)
4th quintile(0.567-0.671)	4,856	2,676.1	10	3.74(2.01, 6.95)
5th quintile(0.671-1.019)	4,856	2,751.8	11	3.99(2.21, 7.22)
Paddy				
1st quintile(0.148-0.279)	4,856	2,759.4	16	5.80(3.55 9.47)
2nd quintile(0.279-0.325)	4,856	2,649.9	15	5.66(3.41 9.39)
3rd quintile(0.325-0.381)	4,856	2,696.6	8	2.97(1.48 5.93)
4th quintile(0.381-0.455)	4,856	2,715.4	31	11.42(8.03 16.23)
5th quintile(0.455-0.732)	4,856	2,769.4	24	8.67(5.81 12.92)
Dry crop land				
1st quintile(0.128-0.332)	4,856	2,749.9	22	8.00(5.27, 12.15)
2nd quintile(0.332-0.377)	4,856	2,744.4	29	10.57(7.34, 15.21)
3rd quintile(0.377-0.439)	4,856	2,755.5	20	7.26(4.68, 11.25)
4th quintile(0.439-0.566)	4,856	2,729.6	9	3.30(1.72, 6.34)
5th quintile(0.566-0.735)	4,856	2,611.1	14	5.36(3.18, 9.05)
Others				
1st quintile(0.412-0.623)	4,856	2,677.1	14	5.23(3.10, 8.83)
2nd quintile(0.623-0.658)	4,856	2,710.1	12	4.43(2.52, 7.80)
3rd quintile(0.658-0.688)	4,856	2,698.2	10	3.71(1.99, 6.89)
4th quintile(0.688-0.718)	4,856	2,704.8	13	4.81(2.79, 8.28)
5th quintile(0.718-1.171)	4,856	2,800.4	45	16.07(11.99, 21.52)

[1]Angular transformed proportion of land use variables in parentheses
[2]Person years at risk
[3]95% confidence interval in parentheses

5.5.8 Results of Poisson regression analysis

P. vivax malaria (dependent variable) together with demographic, human behaviour and environmental variables (independent variables) in the study cohort were fitted into a Poisson regression model. To appraise the associations between vivax malaria and all variables collected in the present study, separate analyses were carried out by fitting each individual variable at a time into a Poisson regression model for the crude analysis as shown in the fifth column of Tables 5.16 and 5.17. The final model was selected by adding new/one variable at a time into the model. Only the variables that were significant, after adjusting for other covariates, were included in the final model. Consequently, ages, ownership of mosquito net, history of temporary migration, ethnic group, altitude, paddy and forest variables were included in the final model. The final model for *P. vivax* malaria was shown as **bold** in the last column of Tables 5.16 and 5.17. The test for trend for each variable was also carried out as shown (*italic P-value*) in the Tables.

The results of the Poisson regression modelling for the risk of *P. vivax* on socio-economic and demographic variables are shown in Table 5.16. The crude analysis shows that the risk of *P. vivax* in females was roughly the same as males. After adjusting for other potential confounding variables in the final model, the relative risk between females and males remains the roughly same. There was a clear association between age group and the risk of *P. vivax* as shown in Table 5.16. The school-age-children were at the highest risk in the crude analysis, the same trend remains in the final model. The younger the age (above 7 years), the higher the risk of *P. vivax* malaria (*P<0.008*), suggesting the young age group lacked immunity to the parasite.

There were large variations in the risk of *P. vivax* malaria among different ethnic groups in the study cohort. The Dai ethnic group was at the highest risk, followed by Miao, Zhuang and Hani groups. Yi ethnic group was at the lowest risk in the crude analysis. The association between Hani and Dai ethnic groups and risk of vivax malaria were not significant in the final model, suggesting the associations identified in the crude analysis are due to the confounding effect of other variables, particularly altitude as shown in Table 5.2. Only the Yi ethnic group had a significantly lower relative risk. The biological reason of the lower relative risk for this group is not clear.

The crude analysis suggested that the higher educated population were at a higher risk of *P. vivax* malaria in the study cohort. The effect of education on the risk of *P. vivax* was not statistically significant after adding the variable into the final model although the same trend still remained, possibly due to chance. The residents of temporary huts appeared to experience a higher risk of malaria. But it was not statistically significant in either the crude analysis or the adjusted analysis in the final model.

The crude analysis indicates that the population owning mosquito nets had a slightly higher risk of *P.vivax* incidence than those who didn't own mosquito nets. People owning a net had 0.73 times the risk of vivax malaria as compared with those without mosquito nets in the final model, suggesting that a protective efficacy of mosquito net against vivax malaria of 27% in the study cohort. The results imply that the higher risk of *P. vivax* malaria among mosquito net owners in the crude analysis was due to the effect of other confounding variables, particularly altitude as shown in Table 5.3.

The effect of temporary migration on the risk of vivax malaria is shown in Table 5.16. People with a history of migration during the study period had 2.84 times risk of vivax malaria than that of population without any history of temporary migration in crude analysis. The effect is much stronger (RR=4.97) in the final model. However, this might not be a proper estimate of the effect of migration activity because our indigenous malaria case definition might include some imported cases as discussed in section 5.5.6 of this Chapter.

The association between the risk of *P. vivax* and the environmental determinants are shown in Table 5.17. Altitude is negatively associated with the risk of vivax malaria. People residing in an area with altitude over 800 metres had less than half the risk of those living below 800 metres. The effects were much stronger in the final model as shown in Table 5.17. The significant differences between crude rate ratio and adjusted rate ratio suggested the unbalanced distribution of other *P. vivax* malaria determinants against altitude in the study area. In other words, other major determinants of *P. vivax* malaria are mainly distributed in the higher altitude areas, such as paddy, forest and migration etc.

The crude analysis indicates a negative correlation between the risk of vivax malaria relative to the amount of forest. But a strong positive correlation was identified in the final model. The apparent negative association in the crude analysis is largely due to the positive correlation between altitude and forest, and negative correlation between paddy and forest as shown in Table 5.8. Nevertheless, the effect of forest dramatically increased the risk of vivax malaria at the fourth quintile with no further increasing at fifth quintile, suggesting that mosquitoes need at least a certain amount of forest (threshold) around the environment to increase their density and transmission. Once the mosquito gets enough forest cover and forest edges to maintain their breeding, more forest does not progressively increase the risk of malaria, probably due to a non-linear correlation between the amount of forest and the amount of mosquito breeding sites such as forest edges in the study area.

The crude analysis suggested the possible positive association between risk of vivax malaria and relative amount of paddy around the households. The association between the risk of *P. vivax* and paddy was even stronger in the final model as was shown in Table 5.17. The more paddy fields, the higher relative risk of vivax malaria.

The crude analyses indicated a strong negative correlation between the risk of *P. vivax* malaria and dry cropland in the study cohort. The appearance of the negative association between the two variables disappeared after adding the variable into the final model, suggesting that the apparently strong negative correlation between the relative amount of dry cropland and the risk of vivax malaria was due to other confounding effects. As shown in Table 5.8, the amount of dry crop land was significantly negatively correlated with paddy and forest in the study cohort. More dry crop land around household means less forest and/or paddy, and therefore fewer mosquito breeding sites. There was no apparent effect of land use category "others" in either crude or adjusted analysis of the Poisson regression modelling as shown in Table 5.17.

Table 5.16. Poisson regression modelling for *P. vivax* indigenous malaria and socio-economic, demographic and behaviour variables in the study cohort, Feng Chun Ling

Variables	No. pop.	PYAR[1]	No. vivax	Crude rate ratio	Adjusted rate ratio[2,3]
Sex					
Male	13,020	7,130.9	128	1	1
Female	11,260	6,459.7	114	0.98(0.76 1.26)	1.19(0.92 1.54)
				P=0.90	*P=0.18*
Ages					
1 - 6	2,723	1,584.4	21	1	**1**
7 – 14	4,314	2,520.0	58	1.74(1.05 2.87)	**1.63(0.99 2.69)**
15 – 30	7,630	4,158.0	84	1.52(0.95 2.45)	**1.14(0.70 1.85)**
31 – 44	4,881	2,645.5	49	1.40(0.84 2.33)	**0.99(0.59 1.66)**
>= 45	4,732	2,682.7	30	0.84(0.48 1.47)	**0.75(0.43 1.32)**
				P<0.009	***P<0.008***
Ethnic groups					
Han	10,840	6,042.0	78	1	**1**
Hani	9,586	5,304.0	124	1.81(1.36 2.40)	**1.23(0.88 1.70)**
Yi	2,579	1,500.6	16	0.83(0.48 1.41)	**0.36(0.20 0.66)**
Dai	651	382.4	15	3.04(1.75 5.28))	**0.73(0.34 1.55)**
Miao	485	279.8	7	1.94(0.89 4.20)	**1.10(0.49 2.45)**
Zhuang	139	81.8	2	1.89(0.47 7.71)	**0.49(0.10 1.93)**
				P<0.001	***P<0.001***
Education					
Illiterate	14,044	7,990.1	123	1	1
Elementary	8,790	4,828.2	100	1.35(1.03 1.75)	1.10(0.84 1.44)
Secondary	1,446	772.3	19	1.60(0.99 2.59)	1.27(0.77 2.07)
				P=0.03	*P=0.58*
Net					
No	19,724	11,017.4	187	1	**1**
Yes	4,556	2,573.2	55	1.26(0.92 1.70)	**0.73(0.50 1.02)**
				P=0.13	***P=0.06***
Quality of house					
Temporary hut	839	460.7	12	1	1
Flat	12,540	7,030.6	112	0.61(0.34 1.11)	0.63(0.34 1.14)
Good	10,901	6,099.3	118	0.74(0.41 1.34)	0.73(0.40 1.34)
				P=0.14	*P=0.22*
Migration					
No	20,908	12,315.7	187	1	**1**
Yes	3,372	1,275.0	55	2.84(2.10 3.83)	**4.97(3.56 6.94)**
				P<0.001	***P<0.001***

[1] Person years at risk

[2] Adjusting for other potential confounding variables namely age, ownership of mosquito net, history of temporary migration, ethnic group, altitude, paddy, forest.

[3] The variables in the final model were bold

Table 5.17. Poisson regression modelling for *P. vivax* indigenous malaria, altitude and land use variables in the study cohort, Feng Chun Ling

Variables[1]	No. pop.	PYAR[2]	No. vivax	Crude rate ratio	Adjusted rate ratio[3,4]
Altitude					
<=800	1,784	1,049.1	43	1	**1**
~1200	5,977	3,439.2	70	0.50(0.34 0.73)	**0.06(0.03 0.13)**
~1350	5,198	2,881.5	52	0.44(0.29 0.66)	**0.08(0.04 0.16)**
~1500	7,510	4,133.0	58	0.34(0.23 0.51)	**0.04(0.02 0.10)**
>1500	3,811	2,087.8	19	0.22(0.13 0.38)	**0.01(0.007 0.06)**
				P<0.001	***P<0.001***
Forest					
1st quintile(0.166-0.404)	4,856	2,787.0	73	1	**1**
2nd quintile(0.404-0.507)	4,856	2,701.6	44	0.62(0.43 0.90)	**1.17(0.76 1.80)**
3rd quintile(0.507-0.567)	4,856	2,674.2	39	0.56(0.38 0.82)	**1.22(0.75 1.99)**
4th quintile(0.567-0.671)	4,856	2,676.1	53	0.76(0.53 1.08)	**2.37(1.44 3.89)**
5th quintile(0.671-1.019)	4,856	2,751.8	33	0.46(0.30 0.69)	**2.17(1.22 3.88)**
				P=0.001	***P=0.002***
Paddy					
1st quintile(0.148-0.279)	4,856	2,759.4	35	1	**1**
2nd quintile(0.279-0.325)	4,856	2,649.9	39	1.16(0.73 1.83)	**1.35(0.83 2.18)**
3rd quintile(0.325-0.381)	4,856	2,696.6	32	0.94(0.58 1.51)	**1.63(0.97 2.73)**
4th quintile(0.381-0.455)	4,856	2,715.4	59	1.71(1.13 2.60)	**2.11(1.34 3.31)**
5th quintile(0.455-0.732)	4,856	2,769.4	77	2.19(1.47 3.27)	**2.90(1.87 4.50)**
				P<0.001	***P<0.001***
Dry crop land					
1st quintile(0.128-0.332)	4,856	2,749.9	70	1	1
2nd quintile(0.332-0.377)	4,856	2,744.4	55	0.79(0.55 1.12)	0.84(0.53 1.34)
3rd quintile(0.377-0.439)	4,856	2,755.5	55	0.78(0.55 1.12)	0.85(0.46 1.58)
4th quintile(0.439-0.566)	4,856	2,729.6	31	0.45(0.29 0.68)	0.77(0.30 1.99)
5th quintile(0.566-0.735)	4,856	2,611.1	31	0.47(0.31 0.71)	0.85(0.19 3.83)
				P<0.001	*P=0.89*
Others					
1st quintile(0.412-0.623)	4,856	2,677.1	36	1	1
2nd quintile(0.623-0.658)	4,856	2,710.1	57	1.56(1.03 2.37)	1.17(0.74 1.83)
3rd quintile(0.658-0.688)	4,856	2,698.2	50	1.38(0.90 2.11)	1.00(0.64 1.57)
4th quintile(0.688-0.718)	4,856	2,704.8	40	1.10(0.70 1.72)	0.72(0.43 1.19)
5th quintile(0.718-1.171)	4,856	2,800.4	59	1.57(1.03 2.37)	0.72(0.40 1.28)
				P=0.12	*P=0.13*

[1] Angular transformed proportion of land use variables in parentheses

[2] Person year at risk

[3] Adjusting for other potential confounding variables namely age, ownership of mosquito net, history of migration, ethnic group, altitude, paddy, forest.

[4] The variables in the final model were bold

A Poisson regression model for *P. falciparum* malaria for the study cohort was also built. Due to the relatively small number of *P. falciparum* malaria cases in the present study, it would not be possible to get enough power to obtain statistical significance on most variables in which we are interested. Therefore, the same set of the variables as in the *P. vivax* malaria Poisson regression model were included in the final model *for P. falciparum* on the assumption of the same aetiology of *P. falciparum* as *P. vivax* malaria in the present study. The final model is shown as **bold** in the last column of Tables 5.18 and 5.19. The methodology of the modelling is the same as described as in the context of *P. vivax.*

The results of Poisson regression modelling for the risk of *P. falciparum* on socio-economic and behaviour variables are shown in Table 5.18. The crude analysis indicates that the risk of *P. falciparum* malaria in females was slightly lower than in males. The relative risks of females and males were roughly the same after adjusting other potential confounding variables in the final model. There were some variations in the rate ratios of *P. falciparum* at different ages as shown in Table 5.18. The pre-school children were at the highest risk in the final model, which reflects that pre-school children had little immunity to *P. falciparum* although the association was not statistically significant due to small number of cases under study.

There were large variations in the risk of *P. falciparum* malaria in different ethnic groups. Dai and Zhuang ethnic groups were at the highest risk, followed by the Miao ethnic group, and Yi ethnic group and the Han Chinese at the lowest risk in the crude modelling. Dai, Hani and Zhuang had a significantly higher risk than Han Chinese, Miao and Zhuang group were at the borderline. The association between ethnic group and risk of falciparum malaria was no longer statistically significant in the final model, suggesting the associations identified in the crude analysis were due to the confounding effect of other variables, particularly altitude, since the spatial distribution of ethnic groups was dependent on altitude as in the context of *P. vivax.*

The crude analysis indicated that individuals with higher education were at a higher risk of *P. falciparum.* But after adding the variable in the final model, the relative risk of *P. falciparum* was roughly the same as shown in Table 5.18. The population living in better houses had a lower risk of *P. falciparum* malaria infection. This trend remained after adjusting for other confounding variables in the final model although it was not statistically significant. This might be due to the small number of *P. falciparum* malaria cases in the "good" group.

People owning mosquito nets had a significantly higher risk of *P.falciparum* infection than those who didnot own mosquito nets in the crude model. In the final model, people with nets only had 0.85 times risk of falciparum malaria as compared with those without mosquito nets, but the difference was not statistically significant. This result reveals that the apparent positive connection between the risk of *P. falciparum* and ownership of mosquito nets is due to a confounding effect, particularly of altitude (Table 5.3). Nevertheless, the absence of strong evidence that the ownership of mosquito nets had protective effect on *P. falciparum* malaria in the study cohort was probably due to the relatively small number of malaria cases in the present analysis.

The effect of temporary migration on the risk of falciparum malaria is shown in Table 5.18. Population with experience of migration had around 16 fold the relative risk (RR=15.76) as shown in the final model. The dramatic increase in the risk of *P. falciparum* malaria suggests that most migration occurred from a high altitude to low altitude area as has already been shown in Table 5.5. Nevertheless, this point parameter might over estimate the true effect of temporary migration activity itself on indigenous malaria due to improper indigenous malaria case definition as in the context of *P. vivax*.

The association of the risk of *P. falciparum* with environmental determinants is shown in Table 5.19. Altitude is negatively associated with the risk of falciparum malaria. The people residing in areas above 800 metres had less than one fifth of the risk of those living in below 800 metres. The same trend remains in the final model as shown in Table 5.19.

The crude analysis indicates a negative correlation between risk of falciparum malaria and the amount of forest. But a positive correlation was identified in the final model although it was not statistically significant due to the relatively small number of falciparum malaria cases in this study. The apparent negative association in the crude analysis is largely due to the positive correlation between altitude and forest as already discussed in the analysis of the risk determinants of *P. vivax* malaria in the study cohort.

A possible positive association between the risk of falciparum malaria and the relative amount of paddy in the study cohort was revealed in the crude analysis. The trend of an association between the risk of falciparum malaria and relative amount of paddy around the households remained in the final model, but was not statistically significant. It might be due to chance, but considering the mosquito ecology and malaria biology as well as the results of the effect of paddy on *P. vivax* malaria in this chapter, this lack of significance might be due to the relatively small number of *P. falciparum* cases so that there is not enough power to detect a significant association.

There was a strongly negative correlation between the risk of *P. falciparum* malaria and dry cropland in the study cohort in the crude analysis. This disappeared after adding the variable in the final model by adjusting for other confounding variables, which implies that the apparent association is due to the confounding effect of other variables, as dry cropland and "other" land use features were significantly correlated with altitude and other land use variables as shown in Table 5.8. A positive correlation was seen between the land use variable "others" and the *P. falciparum* malaria in the crude analysis, but the association disappeared after adding it in the final model to adjust other confounding variables in the study cohort.

Table 5.18. Poisson regression modelling for *P. falciparum* indigenous malaria, socio-economic, demographic and behaviour variables in the study cohort, Feng Chun Ling

Variables	No. pop.	PYAR[1]	No. falcip.	Crude rate ratio	Adjusted rate ratio[2,3]
Sex					
Male	13,020	7,130.9	55	1	1
Female	11,260	6,459.7	39	0.78(0.52 1.18)	1.03(0.68 1.57)
				P=0.24	*P=0.87*
Ages					
1 - 6	2,723	1,584.4	11	1	**1**
7 – 14	4,314	2,520.0	13	0.74(0.33 1.66)	**0.64(0.29 1.44)**
15 – 30	7,630	4,158.0	39	1.35(0.69 2.63)	**0.84(0.42 1.68)**
31 – 44	4,881	2,645.5	20	1.09(0.52 2.27)	**0.63(0.30 1.34)**
>= 45	4,732	2,682.7	11	0.59(0.26 1.36)	**0.50(0.22 1.17)**
				P=0.10	***P=0.41***
Ethnic groups					
Han	10,840	6,042.0	36	1	**1**
Hani	9,586	5,304.0	24	1.71(1.02 2.86)	**1.33(0.74 2.40)**
Yi	2,579	1,500.6	6	1.01(0.41 2.46)	**0.66(0.24 1.81)**
Dai	651	382.4	22	14.48(8.12 25.83)	**2.72(0.96 7.78)**
Miao	485	279.8	3	2.70(0.81 8.96)	**1.81(0.50 6.50)**
Zhuang	139	81.8	3	9.23(2.78 30.66)	**2.00(0.98 8.07)**
				P<0.001	***P=0.24***
Education					
Illiterate	14,044	7,990.1	50	1	1
Elementary	8,790	4,828.2	36	1.19 (0.78 1.83)	0.89(0.57 1.37)
Secondary	1,446	772.3	8	1.66 (0.78 3.49)	1.11(0.52 2.37)
				P=0.37	*P=0.78*
Net					
No	19,724	11,017.4	60	1	**1**
Yes	4,556	2,573.2	34	2.43(1.59 3.70)	**0.85(0.51 1.43)**
				P<0.001	***P=0.55***
Quality of house					
Temporary hut	839	460.7	37	1	1
Flat	12,540	7,030.6	52	0.68 (0.27 1.71)	0.60(0.24 1.52)
Good	10,901	6,099.3	5	0.56 (0.22 1.42)	0.58(0.22 1.49)
				P=0.39	*P=0.52*
Migration					
No	20,908	12,315.7	65	1	**1**
Yes	3,372	1,275.0	29	4.31(2.78 6.68)	**15.76(8.76 28.32)**
				P<0.001	***P<0.001***

[1] Angular transformed proportion of land use variables in parentheses

[2] Adjusting for other potential confounding variables namely age, ownership of mosquito net, history of migration, ethnic group, altitude, paddy, forest

[3] The variables in the final model were bold

Table 5.19. Poisson regression modelling for *P. falciparum* indigenous malaria, altitude and land use variables in the study cohort, Feng Chun Ling

Variables[1]	No. pop	PYAR[2]	No. falcip.	Crude rate ratio	Adjusted rate ratio[3]
Altitude					
<=800	1,784	1,049.1	37	1	**1**
~1200	5,977	3,439.2	21	0.17(0.10 0.30)	**0.16(0.04 0.56)**
~1350	5,198	2,881.5	6	0.06(0.03 0.14)	**0.05(0.01 0.19)**
~1500	7,510	4,133.0	19	0.13(0.08 0.23)	**0.12(0.03 0.47)**
>1500	3,811	2,087.8	11	0.15(0.08 0.29)	**0.12(0.02 0.43)**
				P<0.001	***P<0.001***
Forest					
1st quintile(0.166-0.404)	4,856	2,787.0	44	1	**1**
2nd quintile(0.404-0.507)	4,856	2,701.6	15	0.35(0.20 0.63)	**1.27(0.60 2.69)**
3rd quintile(0.507-0.567)	4,856	2,674.2	14	0.33(0.18 0.61)	**1.34(0.58 3.10)**
4th quintile(0.567-0.671)	4,856	2,676.1	10	0.24(0.12 0.47)	**1.03(0.41 2.60)**
5th quintile(0.671-1.019)	4,856	2,751.8	11	0.25(0.13 0.49)	**1.85(0.70 4.91)**
				P<0.001	***P=0.70***
Paddy					
1st quintile(0.148-0.279)	4,856	2,759.4	16	1	**1**
2nd quintile(0.279-0.325)	4,856	2,649.9	15	0.98(0.48 1.97)	**0.79(0.38 1.68)**
3rd quintile(0.325-0.381)	4,856	2,696.6	8	0.51(0.22 1.20)	**0.93(0.37 2.31)**
4th quintile(0.381-0.455)	4,856	2,715.4	31	1.97(1.08 3.60)	**1.44(0.74 2.81)**
5th quintile(0.455-0.732)	4,856	2,769.4	24	1.49(0.79 2.81)	**1.60(0.78 3.31)**
				P=0.005	***P=0.23***
Dry crop land					
1st quintile(0.128-0.332)	4,856	2,749.9	22	1	1
2nd quintile(0.332-0.377)	4,856	2,744.4	29	1.32(0.76 2.80)	1.14(0.52 2.51)
3rd quintile(0.377-0.439)	4,856	2,755.5	20	0.91(0.50 1.66)	0.93(0.33 2.64)
4th quintile(0.439-0.566)	4,856	2,729.6	9	0.41(0.19 0.90)	1.12(0.24 5.18)
5th quintile(0.566-0.735)	4,856	2,611.1	14	0.67(0.34 1.31)	1.60(0.16 16.18)
				P=0.02	*P=0.89*
Others					
1st quintile(0.412-0.623)	4,856	2,677.1	14	1	1
2nd quintile(0.623-0.658)	4,856	2,710.1	12	0.85(0.39 1.83)	0.61(0.26 1.38)
3rd quintile(0.658-0.688)	4,856	2,698.2	10	0.71(0.31 1.60)	0.45(0.19 1.03)
4th quintile(0.688-0.718)	4,856	2,704.8	13	0.92(0.43 1.96)	0.59(0.25 1.37)
5th quintile(0.718-1.171)	4,856	2,800.4	45	3.07(1.69 5.60)	0.78(0.31 1.97)
				P<0.001	*P=0.39*

[1] Angular transformed proportion of land use variables in parentheses

[2] Person year at risk

[3] Adjusting for other potential confounding variables namely age, owner of mosquito net, history of migration, ethnic group, altitude, paddy, forest,

[4] The variables in the final model were bold.

The Poisson regression models for both *P. vivax* and *P. falciparum* malaria described above are built, based on our indigenous malaria case definition of "someone who in the past month had a history of staying in a high malaria risk area (the low altitude areas) 7 days prior to the current febrile episode". This definition might misclassify some imported cases with an incubation period longer than one month into indigenous cases as already discussed in section 5.5.6. Although we have made proper adjustment on the denominator by using person time at risk and further adjustment by including the migration variable in the final model to correct it, this factor might still affect to build an unbiased model. Therefore, further analysis was carried out to analyse the database using the non-migratory population only. We used the same set of variables as described above to build the Poisson regression models for both *P. vivax* and *P. falciparum* malaria, so that the models will be comparable.

The final models for both *P. vivax* and *P. falciparum* malaria for the non-migratory population only are shown in Tables 4.20, 4.21, 4.22 and 4.23 as **bold**. The overall conclusion on comparing the results from the overall database with using the non-migratory population database only is that it makes very little difference except that the negative correlation between altitude is weaker in the non-migrating population database only for the *P. vivax* Poisson regression model, which is understandable on removing the migration "noise" in the model. The new model for *P. vivax* also suggest the relative risk for *P. vivax* malaria from 800 to 1,500 metres was roughly the same and the possibility of malaria transmission over 1,500 metres is very small. Besides, the new model for *P. vivax* malaria also reveals a negative correlation between dry crop land and risk of malaria, but the test for trend for this association is not statistically significant, suggesting it might be due to chance.

The new model for *P. falciparum* malaria built by using the non-migratory population database was little different from the old model described before. Because of few cases under study, most point estimates have wider confidence intervals. Overall results reveal that the regression model that included the 'migrants' successfully took out the migration effect statistically. However, the crude rate ratios for altitude and *P. vivax* in Table 5.17 were much further from adjusted rate ratios than were those for the non-migrant population in Table 5.21.

Table 5.20. Poisson regression modelling for *P. vivax* malaria, socio-economic, demographic and behaviour variables by using non-migratory population database of the study cohort.

Variables	No. pop.	No. vivax	Crude rate ratio	Adjusted rate ratio[1,2]
Sex				
Male	10,500	97	1	1
Female	10,408	90	0.94(0.70 1.24)	0.98(0.74 1.31)
			P=0.62	*P=0.90*
Ages				
1 - 6	2,652	20	1	**1**
7 – 14	4,195	50	1.58(0.94 2.65)	**1.50(0.89 2.53)**
15 – 30	6,050	56	1.23(0.74 2.05)	**1.15(0.69 1.92)**
31 – 44	3,776	35	1.23(0.71 2.13)	**1.10(0.64 1.91)**
>= 45	4,235	26	0.81(0.45 1.46)	**0.81(0.45 1.45)**
			P=0.08	***P=0.13***
Ethnic groups				
Han	9,038	59	1	**1**
Hani	8,129	91	1.71(1.24 2.38)	**1.11(0.76 1.61)**
Yi	2,492	16	0.98(0.57 1.71)	**0.39(0.21 0.72)**
Dai	648	15	3.55(2.01 6.25)	**0.94(0.43 2.04)**
Miao	463	4	1.32(0.48 3.64)	**0.70(0.25 1.96)**
Zhuang	138	2	2.22(0.54 9.07)	**0.75(0.17 3.25)**
			P<0.001	***P=0.012***
Education				
Illiterate	12,600	97	1	1
Elementary	7,211	77	1.39(1.03 1.87)	1.23(0.91 1.68)
Secondary	1,097	13	1.54(0.86 2.75)	1.42(0.79 2.56)
			P=0.06	*P=0.28*
Net				
No	16,867	143	1	**1**
Yes	4,041	44	1.28(0.92 1.80)	**0.74(0.51 1.07)**
			P=0.14	***P=0.11***
Quality of house				
Temporary hut	672	9	1	1
Flat	10,750	85	0.59(0.30 1.17)	0.54(0.27 1.08)
Good	9,486	93	0.73(0.37 1.45)	0.67(0.34 1.34)
			P=0.17	*P=0.12*

[1] Adjusting for other potential confounding variables namely age, ownership of mosquito net, ethnic group, altitude, paddy, forest.
[2] The variables in the final model were bold.

Table 5.21. Poisson regression modelling for *P. vivax* malaria, altitude and land use variables by using non-migratory population database of the study cohort

Variables[1]	No. pop.	No. vivax	Crude rate ratio	Adjusted rate ratio[2,3]
Altitude				
<=800	1,783	43	1	**1**
~1200	5,579	62	0.46(0.31 0.68)	**0.14(0.08 0.25)**
~1350	4,344	28	0.27(0.17 0.43)	**0.15(0.08 0.26)**
~1500	6,140	40	0.27(0.18 0.42)	**0.13(0.07 0.24)**
>1500	3,062	14	0.19(0.10 0.35)	**0.06(0.03 0.14)**
			P<0.001	***P<0.001***
Forest				
1st quintile(0.166-0.404)	4,512	66	1	**1**
2nd quintile(0.404-0.507)	4,217	34	0.55(0.36 0.83)	**1.16(0.71 1.88)**
3rd quintile(0.507-0.567)	3,988	27	0.46(0.30 0.72)	**1.42(0.81 2.49)**
4th quintile(0.567-0.671)	3,959	32	0.55(0.36 0.84)	**2.66(1.45 4.89)**
5th quintile(0.671-1.019)	4,232	28	0.45(0.29 0.70)	**2.24(1.20 4.17)**
			P<0.001	***P=0.01***
Paddy				
1st quintile(0.148-0.279)	4,272	29	1	**1**
2nd quintile(0.279-0.325)	3,877	25	0.95(0.56 1.62)	**1.17(0.68 2.02)**
3rd quintile(0.325-0.381)	4,199	23	0.81(0.47 1.39)	**1.55(0.86 2.81)**
4th quintile(0.381-0.455)	4,086	43	1.55(0.97 2.48)	**1.98(1.21 3.26)**
5th quintile(0.455-0.732)	4,474	67	2.21(1.43 3.41)	**2.74(1.72 4.36)**
			P<0.001	***P<0.001***
Dry crop land				
1st quintile(0.128-0.332)	4,299	62	1	1
2nd quintile(0.332-0.377)	4,321	40	0.64(0.43 0.96)	0.69(0.46 1.05)
3rd quintile(0.377-0.439)	4,285	48	0.78(0.53 1.13)	0.79(0.50 1.25)
4th quintile(0.439-0.566)	4,275	21	0.34(0.21 0.56)	0.54(0.30 0.96)
5th quintile(0.566-0.735)	3,728	16	0.30(0.17 0.52)	0.47(0.24 0.90)
			P<0.001	*P=0.10*
Others				
1st quintile(0.412-0.623)	3,984	31	1	1
2nd quintile(0.623-0.658)	4,188	40	1.23(0.77 1.96)	0.93(0.57 1.52)
3rd quintile(0.658-0.688)	4,014	29	0.93(0.56 1.54)	0.77(0.46 1.30)
4th quintile(0.688-0.718)	4,128	29	0.90(0.54 1.50)	0.68(0.40 1.16)
5th quintile(0.718-1.171)	4,594	58	1.62(1.05 2.51)	0.63(0.35 1.14)
			P=0.003	*P=0.42*

[1] Angular transformed proportion of land use variables in parentheses

[2] Adjusting for other potential confounding variables namely age, ownership of mosquito net, ethnic group, altitude, paddy, forest.

[3] The variables in the final model were bold.

Table 5.22.Poisson regression modelling for *P. falciparum* malaria socio-economic, demographic and behaviour variables by using non-migratory population database of the study cohort

Variables	No. pop.	No. falcip.	Crude rate ratio	Adjusted rate ratio[1,2]
Sex				
Male	10,500	40	1	1
Female	10,408	25	0.63(0.38 1.04)	0.70(0.42 1.15)
			P=0.07	*P=0.15*
Ages				
1 - 6	2,652	11	1	**1**
7 – 14	4,195	10	0.57(0.24 1.35)	**0.50(0.21 1.18)**
15 – 30	6,050	25	1.00(0.49 2.02)	**0.82(0.40 1.67)**
31 – 44	3,776	9	0.57(0.24 1.39)	**0.44(0.18 1.07)**
>= 45	4,235	10	0.57(0.24 1.34)	**0.60(0.25 1.41)**
			P=0.29	***P=0.26***
Ethnic groups				
Han	9,038	12	1	**1**
Hani	8,129	20	1.85(0.91 3.79)	**0.72(0.31 1.69)**
Yi	2,492	5	1.51(0.53 4.29)	**0.43(0.13 1.40)**
Dai	648	22	25.57(12.66 51.67)	**1.45(0.47 4.48)**
Miao	463	3	4.88(1.38 17.29)	**1.51(0.39 5.88)**
Zhuang	138	3	16.37(4.62 58.02)	**1.67(0.40 6.95)**
			P<0.001	***P=0.34***
Education				
Illiterate	12,600	34	1	1
Elementary	7,211	24	1.23(0.73 2.08)	0.99(0.58 1.70)
Secondary	1,097	7	2.36(1.05 5.33)	1.76(0.77 4.04)
			P=0.11	*P=0.38*
Net				
No	16,867	33	1	**1**
Yes	4,041	32	4.05(2.49 6.58)	**1.10(0.62 1.97)**
			P<0.001	***P=0.74***
Quality of house				
Temporary hut	672	3	1	1
Flat	10,750	37	0.77(0.24 2.50)	0.50(0.15 1.64)
Good	9,486	25	0.59(0.18 1.96)	0.51(0.15 1.70)
			P=0.48	*P=0.51*

[1] Adjusting for other potential confounding variables namely age, ownership of mosquito net, ethnic group, altitude, paddy, forest.
[2] The variables in the final model were bold

Table 5.23. Poisson regression modelling for *P. falciparum* malaria, altitude and land use variables by using non-migratory population database of the study cohort

Variables[1]	No. pop.	No. falcip.	Crude rate ratio	Adjusted rate ratio[2,3]
Altitude				
<=800	1,783	37	1	**1**
~1200	5,579	17	0.15(0.08 0.26)	**0.14(0.04 0.42)**
~1350	4,344	2	0.02(0.005 0.09)	**0.04(0.007 0.17)**
~1500	6,140	7	0.05(0.02 0.12)	**0.10(0.033 0.32)**
>1500	3,062	2	0.03(0.008 0.13)	**0.05(0.008 0.33)**
			P<0.001	***P<0.001***
Forest				
1st quintile(0.166-0.404)	4,512	43	1	**1**
2nd quintile(0.404-0.507)	4,217	8	0.20(0.09 0.42)	**1.25(0.47 3.36)**
3rd quintile(0.507-0.567)	3,988	5	0.13(0.052 0.33)	**1.40(0.42 4.71)**
4th quintile(0.567-0.671)	3,959	3	0.08(0.02 0.26)	**1.38(0.29 6.57)**
5th quintile(0.671-1.019)	4,232	6	0.15(0.06 0.35)	**2.64(0.67 10.33)**
			P<0.001	***P=0.21***
Paddy				
1st quintile(0.148-0.279)	4,272	11	1	**1**
2nd quintile(0.279-0.325)	3,877	8	0.80(0.32 1.99)	**0.75(0.30 1.90)**
3rd quintile(0.325-0.381)	4,199	2	0.18(0.04 0.83)	**0.46(0.09 2.21)**
4th quintile(0.381-0.455)	4,086	27	2.57(1.27 5.17)	**1.80(0.80 4.04)**
5th quintile(0.455-0.732)	4,474	17	1.48(0.69 3.15)	**1.45(0.60 3.47)**
			P<0.001	***P=0.15***
Dry crop land				
1st quintile(0.128-0.332)	4,299	16	1	1
2nd quintile(0.332-0.377)	4,321	23	1.43(0.76 2.71)	1.00(0.51 1.98)
3rd quintile(0.377-0.439)	4,285	17	1.07(0.54 2.11)	0.86(0.40 1.88)
4th quintile(0.439-0.566)	4,275	5	0.31(0.12 0.86)	1.12(0.34 3.69)
5th quintile(0.566-0.735)	3,728	4	0.29(0.10 0.86)	1.38(0.35 5.43)
			P=0.003	*P=0.96*
Others				
1st quintile(0.412-0.623)	3,984	6	1	1
2nd quintile(0.623-0.658)	4,188	8	1.27(0.44 3.66)	0.67(0.22 2.07)
3rd quintile(0.658-0.688)	4,014	3	0.50(0.12 1.98)	0.29(0.07 1.18)
4th quintile(0.688-0.718)	4,128	7	1.13(0.38 3.35)	0.63(0.20 2.03)
5th quintile(0.718-1.171)	4,594	41	5.93(2.52 13.96)	0.37(0.09 1.48)
			P<0.001	*P=0.43*

[1] Angular transformed proportion of land use variables in parentheses
[2] Adjusting for other potential confounding variables namely age, ownership of mosquito net, ethnic group, altitude, paddy, forest.
[3] The variables in the final model were bold

5.5.9 Multilevel modelling of Poisson regression

5.5.9.1 Multilevel model of Poisson regression

As the study populations were spatially clustered within the same villages and same households, the risk of malaria for individuals within the same household and same village is likely to be similar. In other words, they were probably spatially auto-correlated within the same households and/or same villages. This spatial autocorrelation is most likely due to their sharing the same socio-economic, landscape and environment factors, which we can model through mathematical or statistical modelling of the variables we have already collected. The spatial autocorrelation might be also due to factors (variables) that we have not collected during the study or that we could not measure properly, such as quality of case detection. The spatial autocorrelation implies that the observations are interdependent in the same households and/or villages and therefore violate the principal assumption of independence in standard statistical approaches to analysis. Ignoring the spatial auto-correlation will give a falsely accurate estimate of parameters and confidence intervals (Goldstein, 1995). Multilevel modelling was used to correct for spatial autocorrelation within the same household and village in the present study (Johnes & Duncan, 1996; Kreft & Leeuw, 1998).

Three levels of approaches were used in this work. Poisson regression was used to model the malaria incidence rates with landscape, environmental, socio-economic and behaviour variables as covariates. λ_{ijk} is the expected rate of malaria in the population i of household j in the village k. The Poisson model can be written in standard log-linear form:

$$\mathrm{Log}_e(\lambda_{ijk}) = \beta_o + \beta X_{ijk} + \nu_k + \mu_{jk} + \epsilon_{ijk}$$

$\mathrm{Log}_e(\lambda_{ijk})$ is the sum of a fixed part, $\beta o + \beta X_{ijk}$, and random part, $\nu_k + \mu_{jk} + \epsilon_{ijk}$. β is the vector of parameters corresponding to the vector of covariates. ν_k is a random effect with variance of σ^2_{ν}reflecting the variation between villages. μ_{jk} is a random effect with variance of σ^2_{μ}reflecting the variation between households. The level 1 random part ϵ_{ijk} is defined by a dummy variable equal to the square root of the expected count and variance constrained to 1 when assuming a Poisson distribution (Goldstein, 1995 & Goldstein *et al.*, 1998).

The model is estimated by using the *MlwiN* software package (Goldstein *et al.*, 1998). The estimation procedure is iterative with first order maximum quasi-likelihood (Goldstein & Rasbash, 1996). Only those variables that were included in the ordinary Poisson regression analysis were included in the final parsimonious multilevel model. The antilog value of the coefficients of the multilevel Poisson regression is used to estimate the rate ratio (Armitage & Berry, 1987, Clayton & Hill, 1993).

5.5.9.2 Results of multilevel modelling on *P. vivax*

The results of multilevel Poisson regression modelling for the risk of *P. vivax* indigenous malaria with different socio-economic, behaviour and environmental variables are shown in Tables 5.24 and 5.25. The variables included in the final multilevel Poisson regression model are shown in the last column in the Tables 5.24 and 5.25 as **bold**. The results of the final ordinary Poisson regression analysis are shown in the second columns of Tables 5.24 and 5.25 for comparison. As shown in the Tables, there were few differences in estimated rate ratios (RRs) of the variables except for mosquito nets, but in most cases the confidence intervals of RRs in multilevel model were much wider than those in ordinary Poisson regression analysis. The results suggest that the autocorrelation within the same household and villages did not play very important role in most variables, particularly for altitude and land use variables, in contrast to the results of multilevel modelling in the phase I study. The protective efficacy of mosquito nets rises from 27% in the ordinary Poisson regression model to 40% on removing the clustering effect in the multilevel model.

As with the results of ordinary Poisson regression modelling, multilevel analyses also identified a negative correlation between the risk of *P. vivax* and altitude in the final model. The results of multilevel modelling on the effect of forest on the risk of *P.vivax* malaria showed a positive correlation between two variables. The rate ratios of different quintiles of forest in the multilevel modelling were roughly the same as the rate ratios estimated by the ordinary Poisson regression model. It also shows that once the relative amount of forest reached a certain threshold, no further increase occurred in the relative risk of vivax malaria, as discussed in section 5.5.7 of this Chapter. Similarly, a positive association between risk of vivax malaria and relative amount of paddy was also demonstrated in the multilevel model. The more paddy field, the higher risk of vivax malaria in the study area.

The results of the multilevel Poisson regression modelling for *P. vivax* indigenous malaria for all the study cohort also show significant autocorrelations within the same villages and same households as shown in the last row of Table 5.25. The autocorrelation could not be explained by the variables within the final model. The results imply that other characteristics (variables) within households and villages must determine malaria transmission or distribution within the household, which we did not measure and include in the final model. A further study of the residuals of household and village variances might help us to understand this spatial autocorrelation. The results of residual analysis might also help the disease control programme (Diamond *et al.*, 1999). It obviously goes beyond scope of the present thesis work.

Table 5.24. Multilevel Poisson modelling for *P. vivax* indigenous malaria socio-economic, demographic and behaviour variables in the study cohort, Feng Chun Ling

Variables	Adjusted rate ratio[1]	Crude rate ratio[2]	Adjusted rate ratio[3,4]
Sex			
Male	1	1	1
Female	1.19(0.92 1.54)	0.98(0.76 1.26)	1.19(0.92 1.53)
Ages			
1 - 6	**1**	1	**1**
7 – 14	**1.63(0.99 2.69)**	1.69(1.04 2.78)	**1.60(0 .97 2.61)**
15 – 30	**1.14(0.70 1.85)**	1.49(0.92 2.40)	**1.20(0.74 1.94)**
31 – 44	**0.99(0.59 1.66)**	1.37(0.82 2.29)	**0.98(0.59 1.64)**
>= 45	**0.75(0.43 1.32)**	0.85(0.49 1.47)	**0.72(0.44 1.27)**
Ethnic groups			
Han	**1**	1	**1**
Hani	**1.23(0.88 1.70)**	1.68(1.12 2.52)	**1.25(0.85 1.86)**
Yi	**0.36(0.20 0.66)**	0.78(0.40 1.50)	**0.35(0.18 0.67)**
Dai	**0.73(0.34 1.55)**	1.99(0.73 5.33)	**0.60(0.23 1.56)**
Miao	**1.10(0.49 2.45)**	2.02(0.76 5.41)	**1.04(0.41 2.67)**
Zhuang	**0.49(0.10 1.93)**	1.44(0.21 9.99)	**0.39(0.07 2.19)**
Education			
Illiterate	1	1	1
Elementary	1.10(0.84 1.44)	1.21(0.92 1.58)	1.03(0.79 1.36)
Secondary	1.27(0.77 2.07)	1.38(0.79 2.41)	1.17(0.71 1.93)
Net			
No	**1**	1	**1**
Yes	**0.73(0.50 1.02)**	0.75(0 .50 1.14)	**0.60(0.40 0.88)**
Quality of house			
Temporary hut	1	1	1
Flat	0.63(0.34 1.14)	0.60(0.33 1.08)	0.64(0.34 1.19)
Good	0.73(0.40 1.34)	0.63(0.34 1.16)	0.71(0.38 1.33)
Migration			
No	**1**	1	**1**
Yes	**4.97(3.56 6.94)**	4.01(3.03 5.29)	**5.10(3.61 7.20)**

[1]Adjusted analysis in ordinary Poisson regression model
[2]Crude analysis in the multilevel Poisson regression model
[3]Adjusted for ages, owner of mosquito net, history of migration, ethnic group, altitude, paddy in the multilevel Poisson regression model
[4]The variables in the final model were bold

Table 5.25. Multilevel Poisson regression modelling for *P. vivax* indigenous malaria, altitude and land use variables in the study cohort, Feng Chun Ling

Variables[1]	Adjusted rate ratio[2]	Crude rate ratio[3]	Adjusted rate ratio[4,5]
Altitude			
<=800	**1**	1	**1**
~1200	**0.06(0.03 0.13)**	0.51(0.28 0.91)	**0.05(0.02 0.13)**
~1350	**0.08(0.04 0.16)**	0.46(0.24 0.84)	**0.07(0.03 0.16)**
~1500	**0.04(0.02 0.10)**	0.36(0.20 0.64)	**0.04(0.01 0.10)**
>1500	**0.01(0.007 0.06)**	0.20(0.09 0.41)	**0.02(0.005 0.05)**
Forest			
1st quintile(0.166-0.404)	**1**	1	**1**
2nd quintile(0.404-0.507)	**1.17(0.76 1.80)**	0.69(0.44 1.08)	**1.18(0.68 2.06)**
3rd quintile(0.507-0.567)	**1.22(0.75 1.99)**	0.56(0.33 0.93)	**1.16(0.64 2.10)**
4th quintile(0.567-0.671)	**2.37(1.44 3.89)**	0.81(0.49 1.32)	**2.63(1.42 4.87)**
5th quintile(0.671-1.019)	**2.17(1.22 3.88)**	0.47(0.27 0.81)	**2.25(1.02 4.60)**
Paddy			
1st quintile(0.148-0.279)	**1**	1	**1**
2nd quintile(0.279-0.325)	**1.35(0.83 2.18)**	1.01(0.57 1.78)	**1.17(0.66 2.09)**
3rd quintile(0.325-0.381)	**1.63(0.97 2.73)**	0.88(0.49 1.57)	**1.53(0.83 2.83)**
4th quintile(0.381-0.455)	**2.11(1.34 3.31)**	1.67(0.96 2.90)	**2.31(1.30 4.11)**
5th quintile(0.455-0.732)	**2.90(1.87 4.50)**	1.92(1.11 3.32)	**2.71(1.53 4.76)**
Dry crop land			
1st quintile(0.128-0.332)	1	1	1
2nd quintile(0.332-0.377)	0.84(0.53 1.34)	0.77(0.49 1.22)	0.88(0.55 1.40)
3rd quintile(0.377-0.439)	0.85(0.46 1.58)	0.92(0.58 1.47)	1.24(0.71 2.14)
4th quintile(0.439-0.566)	0.77(0.30 1.99)	0.50(0.29 0.87)	1.05(0.54 2.04)
5th quintile(0.566-0.735)	0.85(0.19 3.83)	0.46(0.25 0.84)	1.05(0.46 2.42)
Others			
1st quintile(0.412-0.623)	1	1	1
2nd quintile(0.623-0.658)	1.17(0.74 1.83)	1.54(0.92 2.56)	1.23(0.73 2.01)
3rd quintile(0.658-0.688)	1.00(0.64 1.57)	1.55(0.91 2.62)	1.03(0.61 1.87)
4th quintile(0.688-0.718)	0.72(0.43 1.19)	1.08(0.62 1.92)	0.65(0.36 1.18)
5th quintile(0.718-1.171)	0.72(0.40 1.28)	1.57(0.90 2.74)	0.62(0.31 1.30)
Random effect			
Village level variance(SE)	---	---	**0.15(0.08)**
Household level variance(SE)	---	---	**0.78(0.32)**
Individual level variance(SE)	---	---	**1**

[1]Angular transformed proportion of land use variables in parentheses
[2]Adjusted analysis with ordinary Poisson regression
[3]Multilevel modelling controlling for spatial autocorrelation of households and villages
[4]Multilevel modelling controlling spatial autocorrelation and other confounding factors, ages, owner of mosquito net, history of migration, ethnic groups, altitude, paddy, forest,
[5]The variables in the final model were bold

5.5.9.3 The results of multilevel modelling on *P. falciparum*

A similar method was used to analyse *P. falciparum* indigenous malaria with the landscape, environmental and socio-economic variables. The results of multilevel Poisson regression modelling for the relative risk of *P. falciparum* on demographic and human behaviour variables are shown in Tables 5.26 and 5.27. The results of ordinary Poisson regression modelling are also shown in the second columns of the tables for comparison. The variables included in the final Poisson regression model are shown in the last column of Tables 5.26 and 5.27 as **bold.** The overall conclusion on comparing the results in the ordinary Poisson regression model with those brought in by the multilevel Poisson regression analysis makes very little difference except to the rate ratio in the Dai ethnic group where the rate ratio falls from 2.72 to 1.84 and in relation to the protective efficacy of mosquito nets, where the protective effect rises from 15% to 29% on removing the clustering effect. Migration also has a smaller but still very great effect as the rate ratios falls from 15.76 to 13.74, leaving no doubt about the importance of migration.

Similarly, the multilevel Poisson regression modelling also identifies a negative association of altitude and the risk of falciparum malaria in the final model, but the point estimates of rate ratio's confidence intervals are much wider in the final multilevel model. The results of the multilevel analysis also identified the possible positively correlation between forest and falciparum malaria, and paddy and falciparum malaria as shown in Table 5.27, but the associations are not statistically significant probably due to the small number of cases in the study.

The model also reveals a significant variance within household that could not explained by the final model as shown in the last row of Table 5.27. The interpretation for the variance is the same as in the context of *P. vivax* malaria above.

Table 5.26. Multilevel Poisson regression modelling for *P. falciparum* indigenous malaria, socio-economic, demographic and behaviour variables in the study cohort, Feng Chun Lin

Variables	Adjusted rate ratio[1]	Crude rate ratio[2]	Adjusted rate ratio[3,4]
Sex			
Male	1	1	1
Female	1.03(0.68 1.57)	0.78(0.54 1.14)	1.03(0.67 1.60)
Ages			
1 - 6	**1**	1	**1**
7 – 14	**0.64(0.29 1.44)**	0.66(0.32 1.37)	**0.64(0.27 1.50)**
15 – 30	**0.84(0.42 1.68)**	1.25(0.69 2.27)	**0.90(0.44 1.84)**
31 – 44	**0.63(0.30 1.34)**	0.99(0.51 1.91)	**0.65(0.30 1.43)**
>= 45	**0.50(0.22 1.17)**	0.59(0.28 1.23)	**0.49(0.22 1.22)**
Ethnic groups			
Han	**1**	1	**1**
Hani	**1.33(0.74 2.40)**	1.58(0.83 2.99)	**1.40(0.72 2.71)**
Yi	**0.66(0.24 1.81)**	0.86(0.30 2.42)	**0.64(0.21 1.95)**
Dai	**2.72(0.96 7.78)**	6.07(1.75 20.98)	**1.84(0.48 7.05)**
Miao	**1.81(0.50 6.50)**	2.98(0.74 12.06)	**1.73(0.38 2.06)**
Zhuang	**2.00(0.98 8.07)**	8.06(1.10 58.93)	**2.26(0.34 14.40)**
Education			
Illiterate	1	1	1
Elementary	0.89(0.57 1.37)	1.06(0.71 1.59)	0.92(0.58 1.41)
Secondary	1.11(0.52 2.37)	1.47(0.75 2.87)	1.21(0.56 2.64)
Net			
No	**1**	1	**1**
Yes	**0.85(0.51 1.43)**	1.06(0.59 1.92)	**0.71(0.38 1.32)**
Quality of house			
Temporary hut	1	1	1
Flat	0.60(0.24 1.52)	0.69(0.28 1.67)	0.62(0.22 1.72)
Good	0.58(0.22 1.49)	0.68(0.26 1.77)	0.63(0.22 1.83)
Migration			
No	**1**	1	**1**
Yes	**15.76(8.76 28.32)**	7.95(5.52 11.45)	**13.74(7.47 25.27)**

[1]Adjusted analysis with ordinary Poisson regression
[2]Multilevel modelling controlling for spatial autocorrelation of households and villages
[3]Multilevel modelling controlling spatial autocorrelation and other confounding factors, ages, owner of mosquito net, history of migration, ethnic groups, altitude, paddy, forest,
[4]The variables in the final model were bold

Table 5.27. Multilevel Poisson regression modelling for *P. falciparum* indigenous malaria, altitude and land use variables in the study cohort, Feng Chun Ling

Variables[1]	Adjusted rate ratio[2]	Crude rate ratio[3]	Adjusted rate ratio[4,5]
Altitude			
<=800	**1**	1	**1**
~1200	**0.16(0.04 0.56)**	0.22(0.10 0.50)	**0.17(0.04 0.80)**
~1350	**0.05(0.01 0.19)**	0.07(0.02 0.21)	**0.05(0.01 0.23)**
~1500	**0.12(0.03 0.47)**	0.17(0.08 0.38)	**0.11(0.02 0.57)**
>1500	**0.12(0.02 0.43)**	0.20(0.08 0.52)	**0.12(0.02 0.82)**
Forest			
1st quintile(0.166-0.404)	**1**	1	**1**
2nd quintile(0.404-0.507)	**1.27(0.60 2.69)**	0.55(0.29 1.06)	**1.46(0.62 3.46)**
3rd quintile(0.507-0.567)	**1.34(0.58 3.10)**	0.42(0.20 0.92)	**1.40(0.53 3.70)**
4th quintile(0.567-0.671)	**1.03(0.41 2.60)**	0.34(0.14 0.77)	**1.19(0.40 3.49)**
5th quintile(0.671-1.019)	**1.85(0.70 4.91)**	0.27(0.11 0.66)	**1.75(0.55 5.57)**
Paddy			
1st quintile(0.148-0.279)	**1**	1	**1**
2nd quintile(0.279-0.325)	**0.79(0.38 1.68)**	0.97(0.41 2.32)	**0.85(0.36 1.97)**
3rd quintile(0.325-0.381)	**0.93(0.37 2.31)**	0.60(0.23 1.59)	**0.89(0.33 2.40)**
4th quintile(0.381-0.455)	**1.44(0.74 2.81)**	1.45(0.58 3.62)	**1.19(0.50 2.83)**
5th quintile(0.455-0.732)	**1.60(0.78 3.31)**	1.17(0.45 3.05)	**1.55(0.66 3.59)**
Dry crop land			
1st quintile(0.128-0.332)	1	1	1
2nd quintile(0.332-0.377)	1.14(0.52 2.51)	0.99(0.47 2.13)	0.90(0.43 1.91)
3rd quintile(0.377-0.439)	0.93(0.33 2.64)	0.99(0.44 2.26)	0.96(0.40 2.34)
4th quintile(0.439-0.566)	1.12(0.24 5.18)	0.69(0.27 1.75)	1.01(0.35 2.95)
5th quintile(0.566-0.735)	1.60(0.16 16.18)	0.99(0.36 2.68)	1.29(0.37 4.42)
Others			
1st quintile(0.412-0.623)	1	1	1
2nd quintile(0.623-0.658)	0.61(0.26 1.38)	0.77(0.32 1.85)	0.66(0.27 1.61)
3rd quintile(0.658-0.688)	0.45(0.19 1.03)	0.79(0.32 1.96)	0.50(0.20 1.25)
4th quintile(0.688-0.718)	0.59(0.25 1.37)	1.06(0.44 2.53)	0.59(0.22 1.55)
5th quintile(0.718-1.171)	0.78(0.31 1.97)	2.84(1.27 6.34)	0.81(0.28 2.38)
Random effect			
Village level variance(SE)	---	---	**0.37(0.22)**
Household level variance(SE)	---	---	**2.85(0.85)**
Individual level variance(SE)	---	---	**1**

[1]Angular transformed proportion of land use variables in parentheses
[2]Adjusted analysis with ordinary Poisson regression
[3]Multilevel modelling controlling for spatial autocorrelation of households and villages
[4]Multilevel modelling controlling spatial autocorrelation and other confounding factors, ages, owner of mosquito net, history of migration, ethnic groups, altitude, paddy, forest,
[5]The variables in the final model were bold

5.6 Discussion

5.6.1 Overview of results

The prospective follow-up study was carried out in Feng Chun Ling Township, Yuanyang County, Yunnan from May 1998 to December1998. Only local permanent residents were included in the study. Overall 96.9% of the population in the study area were recruited into the study cohort. 24,280 people in 5,007 households in 85 village were followed-up for 7 months. Weekly active household to household surveillance was carried out to take blood slides from whoever had had a febrile episode in the past week, to detect malaria parasites. All village and town doctors who practised medicine in Feng Chun Ling formed a passive surveillance system to take blood slides from all who consulted them with a febrile episode during the study period. A total of 5,360 slides were collected during the study period, of which 663 slides were malaria positive. Recurrences and non-indigenous cases were excluded from the analysis. Consequently, 16 recurrent cases and 315 non-indigenous cases were excluded. Of the 334 indigenous cases, 240 cases were due to *P. vivax* infection. 92 cases were due to *P. falciparum,* and 2 cases were mixed infection of both *P. vivax* and *P. falciparum*. The overall risk of *P. vivax* was 17.81 per 1,000 person years at risk and of *P. falciparum* 6.92 per 1,000 person years at risk in the study area.

The geographical positions of all households were detected by GPS in the study area during the study. The co-ordinates of all households were input into GIS software ArcView and linked with other demographic and socio-economic characteristics of the study population as well as malaria data with a common code number. A SPOT 4 image was classified into 4 categories of land use types, paddy, forest, dry cropland and "others" miscellaneous land use types. Overall accuracy of satellite image classification is 84.2%. By overlaying household position coverage with the new land use map derived from the SPOT 4 image in GIS, the land use pattern of all households was calculated. The mean altitudes of all households were obtained by overlaying the household position file with the DEM of the study area derived from topographical and terrain maps in GIS.

We tabulated *P. vivax and P. falciparum* indigenous malaria for different categories of demographic and socio-economic variables, and environmental variables to appraise the general trend of malaria in relation to them. The risks of malaria associated with the different categories of variables were further quantified by Poisson regression modelling. The results indicate that the risk of malaria for males and females was roughly the same. *P. vivax* and *P. falciparum* affect all age groups, but the pre-school, school age children and young adults were at the highest risk of malaria infection. There were very large variations in the risk of malaria in different ethnic groups, but after controlling for potential confounding variables in the final model, most of them are not statistically significant, suggesting the variations in the risk of malaria are due to other confounding factors. Education makes little difference to the risk of either *P. vivax* or *P. falciparum* infection in the present study. There is an inverse relation between house quality and risk of *P.vivax* and *P. falciparum* malaria, but it does not attain statistical significance. Mosquito nets are protective against malaria. The protective efficacy for *P. vivax* is 27% and *P. falciparum* 15% in the ordinary Poisson regression modelling, but it is not statistically significant for *P. falciparum* malaria.

The results of ordinary Poisson regression modelling also revealed that malaria is negatively correlated with the altitude, but positively correlated with the amount of paddy and forest for *P. vivax* malaria. For *P. falciparum*, the same relationships were seen but they did not attain statistical significance for the amount of paddy and forest probably due to the small number of cases under study.

Further Poisson regression models for both *P. vivax* and *P. falciparum* were built by using the non-migratory population only in the study cohort in hope of removing migration "noise" from the model building. The overall conclusion on comparing the results in the overall database with those using the non-migratory population database was that it makes very little difference except that the negative correlation between altitude was weaker in the non-migratory population database. The new model for *P. vivax* suggests that the possibility of malaria transmission over 1,500 metres is very small.

Further analysis was carried out to correct the spatial autocorrelation of indigenous malaria by multilevel modelling of the Poisson regression. The results of multilevel modelling are very similar in the general trend of association of the risk of indigenous malaria and its demographic, socio-economic and environmental indicators. The parameters estimated by multilevel modelling, however, are more conservative, and the confidence intervals of the parameters are much wider as compared with ordinary Poisson regression in most cases.

The results of multilevel Poisson regression modelling for the risk of *P. vivax* indigenous malaria showed that there were few differences in estimated RRs of the variables, but in most cases the confidence intervals of RRs in the multilevel model were much wider than those of ordinary Poisson regression analysis. The protective efficacy of mosquito nets rise from 27% in ordinary Poisson regression model to 40% on adjusting for the clustering effect in multilevel model. There was a strong negative association between *P. vivax* malaria and altitudes. The significant positive association between *P. vivax* malaria and paddy, forest was also identified in the multilevel Poisson regression model.

Similarly, for *P. falciparum,* the overall conclusion on comparing the results in the ordinary Poisson regression model with those bringing in the multilevel Poisson regression modelling makes very little difference except to the rate ratio in the Dai ethnic group where the rate ratio falls from 2.71 to 1.84, and in relation to the protective efficacy of mosquito nets, where its value rises from 15% to 29% on removing the clustering effect. A negative association between risk of *P. falciparum* malaria and altitude was also identified in the multilevel model. Although there are general trends of the more paddy and more forest, the higher risk of *P. falciparum* malaria, they are not statistically significant, probably due to the small number of cases under study.

A crude analysis comparing the risk of malaria among people with a history of temporary migration with those without any history of migration was also carried out. The overall risks for the migratory population and non-migratory population were 118.6 and 12.1 per 100 persons over 7 month study period, implying that the temporary migratory population had around a 10-fold risk greater than that of the non-migratory population during the study period.

5.6.2 Strengths and constrains of study design, data collection and analysis

5.6.2.1 Prospective cohort study

This study is a prospective cohort study. It was designed to avoid methodological flaws such as lacking the accurate spatial distribution of the exposed population and the malaria cases, confounding variable collection, and being aware of the "ecological fallacy" at the stage of analysis. In essence the study was designed to minimise bias wherever possible. The advantage of the cohort study over other observational studies such as a case-control study is that incidence rates for different categories of exposure could be calculated. Therefore, one could calculate the relative risk of malaria over different categories of determinants. The malaria incidences occurred after the exposures had been identified, therefore the socio-economic and behaviour determinants of malaria were not subject to bias by diagnosis. Information on potential confounding variables was collected and the variables can be used to adjust the model to fit the data. Individual exposure and risk of malaria were identified through proper malaria surveillance systems, and the analysis done at each individual level, hence the "ecological fallacy" was avoided by contrast with limitations of the phase I study.

5.6.2.2 Compliance

3.1% of the population in the study area didn't participant in the study. In theory, if the non-compliant population was exposed to high-risk variables, they might have had more malaria. This would result in underestimating the risk in the whole population. Consequently, it might affect the generalisation of the study results, but would not affect the validity of the study. Relatively fewer females than males failed to enrol in the present study for reasons of non-availability or unwillingness to participate. But the relative number is small, therefore, the non-compliance of small proportion of woman might not affect the validity of the study conclusion.

5.6.2.3 Loss to follow-up

The loss to follow up is one of the major limitations of the prospective cohort study. Although we can make proper adjustment for the denominator, the difficulty raised by persons lost to follow-up is that the probability of the loss may be related to the exposure category and the outcome being measured, or both. To limit the loss to the follow-up, only the permanent residents who currently lived in the village were included in the study. People, who were permanent residents but did not live in the indigenous village during study period, were also excluded from the study. Non-permanent residents were highly mobile. Hence it would be difficult to follow them up. Beside the non-permanent residents are mostly from non-malarious or lower malaria endemic areas with little immunity to malaria; but they usually lived at a lower altitude in the valley to plant banana and other tropical fruits at the highest malaria risk. They also had a different socio-economic background. It is not feasible to study temporary residents. Internal migrants were studied at their main villages of residence.

5.6.2.4 Choice of outcome

Malaria morbidity and mortality are recommended as appropriate outcome measures for malaria epidemiological surveys (Russell *et al.*, 1963; Gilles & Warrell, 1993). Death due to malaria is very rare in East and Southeast Asia, particularly in South China and therefore attempts to measure the effect of environmental factors on mortality would necessarily involved a huge population in a very large area and would consequently be prohibitively expensive. Assessment of the classic symptoms of malaria might also be used. This classical picture is, however, altered by many other factors such as acquisition of immunity or use of prophylactic drugs. Therefore, it is a very imprecise means to estimate the environmental effect on the risk of malaria. Serology might be useful to assess the degree of endemicity, especially in areas of malaria endemicity such as present study area (Gilles & Warrell, 1993). The cumulative of seropositivity over accumulated infection makes it especially difficult to link serum positive samples with clinical episodes of malaria. The malaria incidence rate has been central in evaluating the malaria situation and in planning for control of the diseases in China as well as other area in the Southeast Asia and Pacific Region (Tang *et al.*, 1991; Sharma & Kondrashin, 1991). Thus, the selection of this measure was relevant to the control practice in the region and the result should enable some degree of generalisation to other similar areas in the region. Also it is the variable supposedly reported routinely in the data used in phase I. Consequently, malaria morbidity alone was chosen to assess the effect of environmental factors on the risk of malaria in the present study.

5.6.2.5 Definition of malaria

The malaria transmission was considerably lower in the study area as compared with that seen in the highly endemic areas such as Africa. The overall incidence of malaria is 47.75 per thousand persons per year at risk in the study area in the main transmission season. Malaria is unstable and the general population shows a lack of immunity in the study area as in many other areas of Southeast Asia. While symptoms of malaria may begin before parasitemia is microscopically detectable, the presence of parasitemia, even at a very low level, usually implies clinical malaria in this region (WHO, 1988) and most cases with malaria infection are likely to experience fever. Since there are no clinical parameters which alone would permit a reliable diagnosis of malaria, light microscopic detection of malaria parasite appears to be the most reliable method for diagnosis of malaria as a cause of illness (WHO, 1988). Therefore, fever with malaria parasitemia was used as a feasible and sensitive definition of malaria in the study.

Nevertheless, the transmission of malaria is not uniform throughout the study area. The transmission in the lower altitude areas is much higher than that at the top of the mountains. Therefore, the population at the bottom of mountains might obtain more immunity to the parasites leading to a higher threshold of malaria parasitaemia before onset of fever in the lower altitude areas. Therefore, the present malaria definition might not completely reflect the level of transmission but risk of clinical malaria cases.

5.6.2.6 Incubation period

In malariology, the incubation period of malaria is related to the prepatent period. The prepatent period is the minimal time elapsing between the initial sporozoite injection and the first appearance of parasites in the erythrocytes (Gilles & Warrell, 1993). The incubation period is the time elapsing between the initial malaria infection in the human being and the first clinical manifestation (WHO, 1963). The latter is usually at least 2 days longer than the prepatent period. The incubation period is dependent on the type of infection, the strain of parasite, the number of sporozoites injected, immune status of the host and route of infections (Harinasuta & Bunnag, 1988; Glynn, 1994; Glynn & Bradley, 1995a; Glynn & Bradley, 1995b). The shortest incubation periods of *P. falciparum* and *P.vivax* are 7 and 9 days (Coatney *et al*, 1950; 1971; Garnham 1980). In general it tends to be in the range of 9-30 days (Harinasuta & Bunnag, 1988), but mostly about 14-15 days (Gilles & Warrell, 1993). For *P. vivax* it is extremely varied and can be more than 300 days (Garnham, 1980). Knowledge of the prepatent period and incubation period of malaria in the present study helps us to set the criteria for identifying a case as being indigenous or imported one. The definition of a non-indigenous malaria case is someone who in the past month had a history of staying in a high malaria risk area (the low altitude areas) 7 days prior to the current febrile episode in the present study. Nevertheless, we were not aware of the complexity of the problems of temporary migration in the study area at the stage of the study design. We made the criteria during the fieldwork. The criteria were not strict in terms of classifying the sources of infection and might misclassify some malaria cases with a long incubation period (more than one month) into indigenous malaria cases. It was also sensitive to patient 'recalls' error of the date of their visiting the lower altitude areas, which might also lead us to misclassify some imported malaria cases as indigenous cases as discussed in section 5.5.6 of this Chapter. Completely deleting the temporarily migratory population from analysis might help to remove the definition error, but it would have danger of eliminating some indigenous cases and create a selection bias for the study. The temporary migrants had some different characteristics as compared with the general population. Nevertheless, we did carry out the analyses with the database eliminating the migratory population by using ordinary Poisson regression model as shown in Tables 5.24, 5.25, 5.26 and 5.27.

5.6.2.7 Case detection

The field workers visited all individuals weekly in the study cohort to take blood slides from those who had a febrile episode in the past 7 days. The weekly active malaria surveillance system might be subject to loss of some malaria cases in the present study. Snow *et al.*, (1989) found in The Gambia that weekly surveillance detected 74% cases (as compared with daily surveillance on the assumption that 100% were detected on a daily basis, and 25% by the monthly surveillance. Therefore, Snow and colleagues recommended that, where the prevalence of malaria is similar to that in The Gambia, weekly surveillance is the most cost-effective method to detect malaria. Both active and passive surveillance systems were set up simultaneously in the study. So if febrile patients saw their doctors (part of possive surveillance system), the doctors would take a blood slide for malaria parasite examination and report it even though the patient was missed by the active surveillance system. The relatively inconvenient transportation made it very unlikely that patients would see doctors outside the study area. Therefore, the possibility of loss of detection of malaria cases was very small, unless they were completely asymptomatic (as outside our case definition).

5.6.2.8 Data analysis

The statistical analysis of data derived from geographical information systems and remote sensing were complicated by spatial autocorrelation. In the case of malaria, the risk of malaria in the same villages and household tends to be similar due to people infecting each other or because of other variables that we didn't or couldn't measure. There are several ways to correct the spatial autocorrelation. Most geostatisticians will use kriging (Kitanidis, 1997). But kriging can only model one variable with distance as a covariate between the points. Co-kriging allows for modelling more variables (Burrough & McDonnell, 1998). It would be very unstable for modelling more variables simultaneously (Jennifer L. Dungan, personnel communication). Malaria is a multi-factor disease. Therefore, multilevel modelling was used to correct for the spatial autocorrelation within the same household and villages in the present study (Johnes & Duncan, 1996; Kreft & Leeuw, 1998)

5.6.3 Biases

5.6.3.1 GPS errors

It is well known that a single hand receiver of Global Positioning Systems (GPS) is subjected to error. The largest component of the error, selective availability (SA) is an intentional error component added for security purposes at each satellite by the US military up till May of 2000. Apart from SA, the computations of GPS position co-ordinates are subject to error from several other uncontrolled factors: clock errors, atmospheric conditions, GPS receiver noise and the reflectance of satellite signals. The distorted position error ranges from 30 to 100 metres with a single hand GPS receiver (Letham, 1996). To correct the errors, one might use differential global positioning system (DGPS) as described by Hightower *et al.* (1998) which generally reduces the total error budget to 1-5 metres. The DGPS needs two GPS units simultaneously reading from the same set of satellites (one GPS receiver being located at a known geographical co-ordinate point or base and the other being around the field). Then computer software can be used to correct the error and work out accurate geographical positions. The present study area is located in a mountainous area where it is very difficult to identify a known co-ordinate point. Even if a known co-ordinate point could be identified, it would be very difficult to track the same set of satellites for the base point and the locations of the households around the mountainous areas. Therefore, only two single hand GPS receivers were used to detect the geographical position with latitude and longitude co-ordinates in the present study. There is a 30-100 metres household position uncertainty. The error may make the environmental variables of the households in the study area appear more similar. Nevertheless, logistic reasons (the cost of DGPS) did not allow us to use DGPS to obtain more accurate household position data. In present study, we used a 5-minute mean co-ordinate to represent a household position, which might further decrease this uncertainty (Thomas & Lindsay, 2000). It should be noted that the GPS error has no effect on the detection of malaria cases in the study area. Consequently, the effect of the error results in diluting the association between the risk of malaria and their landscape and environmental indicators rather than the direction of this association.

5.6.3.2 Selection bias

Selection of study subjects was made on the basis of set of inclusion and exclusion criteria. Local permanent residents were included in the present study. Temporary populations residing in low valley in the low altitude areas but not permanent resident in the township, who were at great risk of malaria due to relatively lack of immunity, were not included in the study population. Consequently, it would underestimate the total number of cases in the lower altitude areas. Although this underestimates malaria risk, the uncertainty is where the mobile population contracted malaria and other sources of error. Since this temporary population was very mobile, it was very difficult to follow them to get accurate malaria data in the field. To obtain valid results, only local residents were included in the study. The underestimation of malaria risk in the lower altitude area might only dilute the association between altitude and the risk of malaria, it would not affect the direction of the association.

5.6.3.3. Information Bias

Information bias might be introduced during the follow-up study. Fieldworkers might work harder to collect blood slides in the higher risk villages such as low altitude areas. Laboratory technicians might spend more time reading slides that came from the higher risk area or the slides who were suspected to be malaria parasite positive. Strict guidance was given to fieldworkers on how to do surveillance during the follow-up study. Besides, field supervisors, the field co-ordinator and PI randomly visited the fieldworkers and checked their work to ensure proper surveillance had been done in their target villages based on the field regulations. All slides were coded before they were read. Laboratory technicians were blinded as to where the slides came from. Furthermore, all positive blood slides and 10% of negative blood slides were re-examined to ensure the quality of blood slide examination. Therefore, information bias due to collection of blood slides and parasite identification by reading the blood slides would be kept to a minimum.

In summary, the phase II study has indicated that the landscape and environmental, and some socio-economic and behaviour variables are important determinants of malaria spatial distribution in the study area. Although some biases existed, they only diluted the association of the determinants and the risk of malaria. They would not affect the direction of association. The phase II study has proved the following original hypotheses: (1) the population living at a lower altitude has a higher risk of malaria than those in higher altitude areas. (2) people living in areas with more rice fields have a higher risk of infection with malaria; (3) people living near to or in forest areas around their households tend to have higher risk of *P. vivax* and *P. falciparum* infection. The phase II study rejects another of the original hypotheses; that residence in dry crop field areas is protective. It has also uncovered the importance of internal migratory behaviour as a risk factor.

malaria; (3) people living near to or in forest areas around their households tend to have higher risk of *P. vivax* and *P. falciparum* infection. The phase II study rejects another of the original hypotheses; that residence in dry crop field areas is protective. It has also uncovered the importance of internal migratory behaviour as a risk factor.

Chapter 6

General discussion and conclusions

Geographical information systems, integrated with remote sensing, other data sources, malaria routine surveillance data and a cohort study data set were used to map and quantify landscape and environmental features related to malaria spatial distribution in the Red River basin Yuanyang County, Yunnan, China in general, in Feng Chun Ling Township in particular. The attempts of this study were to identify the environmental, socio-economic and human behaviour indicators of malaria, quantify the effect of landscape and environmental variables in space on the risk of malaria, and provide reliable spatial prediction of malaria distribution and outbreaks in the Red River basin area, Yunnan. Here are the general discussion and conclusions on the most pertinent findings in phase I and phase II studies

6.1 The overall strengths and constrains of the strategy of study

The original intention of the study (phase I) was to compare the various potential landscape and environmental determinants with ground surveillance data based on routinely reported clinical malaria data. This proved to have limitations such as the variable quality of reported malaria incidence as discussed in Chapter 4. It was therefore decided to undertake a field survey of one township in the study county, with a population of 24,000 and a relatively high malaria incidence (phase II). A larger population would have been preferable in the phase II study, but that study was at the absolute upper limit of what was logistically and financially feasible.

Ideally, the spatial predictive model developed in the phase II study should be tested by other field surveys, which should be located in the different places representative of the landscape and land use of much of the Red River basin area. If consistent results between the sites were obtained, we could extrapolate the spatial predictive model from the phase 2 study to the whole Red River basin area, and map malaria risk in the Red River basin area based on the model. Logistic reasons and time constraints made it impossible to select other sites to test the model in the present work. Nevertheless, the study area for phase II study was chosen to make it relatively representative of the rest of the Red River basin area, and the study was designed in such a way as to eliminate as many sources of bias as possible, as discussed in section 5.6 in Chapter 5. The results were consistent between the phase 1 and the phase 2 studies regarding the landscape and environmental determinants of malaria spatial distribution. The biologically plausible landscape and land use determinants in the terms of malaria biology and mosquito ecology strengthen our confidence that the spatial predictive model is unlikely to be biased.

Several attempts to use GIS and high resolution remote sensing data to study malaria and other vector borne diseases have appeared in recent years. Most studies focused on identifying vector abundance and densities (Wood *et al.*, 1991; Beck *et al.*, 1994, Reborts *et al.*, 1996; Beck *et al.*, 1997, Dister *et al.*, 1997; Rejmankova *et al.*, 1998; Lothrop & Reisen., 1999; Moncayo *et al.*, 2000). Most of the studies, however, were on a small scale. Only a few studies were validated by repeated studies in the same area (Wood *et al.*, 1991; Wood *et al.*, 1992) or tested in different places of the same region (Beck *et al.*, 1994; Beck *et al.*, 1997). Recent studies have concentrated more on using GIS and low resolution remote sensing data to identify the risk areas for diseases spatially and temporarily (Hay *et al.*, 1998; Thomson *et al.*, 1999; Snow *et al.*, 1999; Lobitz *et al.*, 2000). The difficulty of using GIS and remote sensing to study malaria and other vector borne diseases is the multi-determinant nature of malaria and other vector-borne disease aetiology. The vector presence and suitable temperatures are only two of several criteria that result in malaria and vector borne disease endemic or epidemic occurrence (Gilles & Warrell, 1993; Molineaux, 1988). The relationship between mosquito abundance and malaria transmission varies with vector species and human behaviour. For instance, a study in The Gambia showed a negative correlation between the abundance of mosquitoes and prevalence of malaria due to wide spread use of mosquito nets in the high risk mosquito biting area (Thomson *et al.*, 1994). Similar studies conducted in two areas in Burkina Faso by Robert and colleagues (1985; 1988) found mosquito density in a rice field area was 4 fold higher than that observed in the Savanna, but the inoculation rate was 5-times lower than in the Savanna. They interpreted that as due to widely used mosquito nets in the rice area, a lower survival rate of *An. gambiae* in the rice fields and zoophilic deviation of the parous females of *An. gambiae* in the rice areas (Robert, 1989). Most previous studies, however, used the abundance of mosquitoes as the surrogate for the risk of malaria and vector borne diseases (Beck *et al.*, 1994; Dister *et al.*, 1997; Moncayo *et al.*, 2000). But, using GIS and remote sensing to study malaria only on the basis of mosquito abundance has a number of limitations as discussed. A more accurate spatial predictive model should incorporate landscape and environmental variables with socio-economic and human behaviour data that have great effect on the occurrence of the diseases. To the author's knowledge, the present study is one of the first larger scale studies to use GIS and high distribution. The biologically plausible landscape and land use determinants in the terms of malaria biology and mosquito ecology strengthen our confidence that the spatial predictive model is unlikely to be biased.

Several attempts to use GIS and high resolution remote sensing data to study malaria and other vector borne diseases have appeared in recent years. Most studies focused on identifying vector abundance and densities (Wood *et al.*, 1991; Beck *et al.*, 1994, Reborts *et al.*, 1996; Beck *et al.*, 1997, Dister *et al.*, 1997; Rejmankova *et al.*, 1998; Lothrop & Reisen., 1999; Moncayo *et al.*, 2000). Most of the studies, however, were on a small scale. Only a few studies were validated by repeated studies in the same area (Wood *et al.*, 1991; Wood *et al.*, 1992) or tested in different places of the same region (Beck *et al.*, 1994; Beck *et al.*, 1997). Recent studies have concentrated more on using GIS and low resolution remote sensing data to identify the risk areas for diseases spatially and temporarily (Hay *et al.*, 1998; Thomson *et al.*, 1999; Snow *et al.*, 1999; Lobtz *et al.*, 2000). The difficulty of using GIS and remote sensing to study malaria and other vector borne diseases is the multi-determinant nature of malaria and other vector-borne disease aetiology. The vector presence and suitable temperatures are only two of several criteria that result in malaria and vector borne disease endemic or epidemic occurrence (Gilles & Warrell, 1993; Molineaux, 1988). The relationship between mosquito abundance and malaria transmission varies with vector species and human behaviour. For instance, a study in The Gambia showed a negative correlation between the abundance of mosquitoes and prevalence of malaria due to wide spread use of mosquito nets in the high risk mosquito biting area (Thomson *et al.*, 1994). Similar studies conducted in two areas in Burkina Faso by Robert and colleagues (1985; 1988) found mosquito density in a rice field area was 4 fold higher than that observed in the Savanna, but the inoculation rate was 5-times lower than in the Savanna. They interpreted that as due to widely used mosquito nets in the rice area, a lower survival rate of *An. gambiae* in the rice fields and zoophilic deviation of the parous females of *An. gambiae* in the rice areas (Robert, 1989). Most previous studies, however, used the abundance of mosquitoes as the surrogate for the risk of malaria and vector borne diseases (Beck *et al.*, 1994; Dister *et al.*, 1997; Moncayo *et al.*, 2000). But, using GIS and remote sensing to study malaria only on the basis of mosquito abundance has a number of limitations as discussed. A more accurate spatial predictive model should incorporate landscape and environmental variables with socio-economic and human behaviour data that have great effect on the occurrence of the diseases. To the author's knowledge, the present study is one of the first larger scale studies to use GIS and high resolution remotely sensed data to study malaria local variation and to directly link malaria and its landscape and land use determinants and human behaviour factors. A spatial prediction model of malaria risk was developed accordingly.

6.2. Findings from phase I and phase II studies

6.2.1 The effect of altitude on the risk of malaria

The effect of altitude on malaria has been discussed in Chapters 4 and 5. The ecological analysis in the phase I study indicated a negative correlation between the mean altitude and the relative risk of malaria for both *P. vivax* and *P. falciparum*. People living in the administrative villages with an altitude over 1,000 metres had less than 13% of the risk of *P. vivax* malaria infection than those who lived in the villages with mean altitude under 1,000 metres in the multilevel logistic regression analysis. For *P. falciparum*, the statistical results showed a similar negative association between the two variables, but the association was not as obvious as that between *P. vivax* and altitude. Those might be due to several reasons. First of all, the number of *P. falciparum* malaria cases was very small in the county during the study period, therefore the results of the analysis would be very unstable. There would be even more difficulty in multilevel modelling analysis as shown in Chapter 4. Secondly, migration would seriously bias the results of the analysis, particularly for *P. falciparum*. People living at the middle and top of mountains went to lower altitude areas to work, and they might be infected with *P. falciparum* during their work in the lower altitude areas. Although we intended to filter away non-indigenous cases by strict criteria of case inclusion and excluded most cases liable to be due to migration, mobile population, non-permanent residents and populations in the towns from study, there was still the possibility of misclassification of the sources of infection among the study population. The misclassification would result in a falsely high risk of malaria in the high altitude areas. Subsequently, it diluted the association between the risk of falciparum malaria and altitude as shown in Table 4.13 in Chapter 4. Nevertheless, less decisive conclusions could be drawn in the phase 1 study due to the problems of reporting variation in different villages and townships, the "ecological fallacy", and unavailability of potentially confounding factors for correction of the models, as discussed in Chapter 4.

The prospective cohort study in the phase II study was carried out in a relatively small area of Yuangyang in Feng Chun Ling Township to further quantify the effect of altitude on the risk of malaria as described in Chapter 5. Both ordinary Possion regression and multilevel Poisson regression was used to model the association between the indigenous malaria and altitude. There is little difference in estimated parameters (RRs) between the two models. The results of the analysis shows the effect of altitude on the risk of malaria was consistent with that in the phase 1 study. But the association between risk of malaria and altitude is much stronger for both *P. vivax* and *P. falciparum* malaria infection in the prospective cohort study than in the ecological analysis in the phase I study, as shown in Tables 5.25 and 5.27. Nevertheless, possible error might enter the models described above due to improper indigenous malaria case definition as discussed in Chapter 5. A further analysis by using all the non-migratory population database was carried out (Tables 5. 21, 5.22, 5.23 and 5.25). The results of the analysis showed that people living in the areas with altitude above 800 metres had only less than 0.15 time risk of malaria as compared with population living below 800 metres. The possibility of malaria transmission was very small in the areas with altitude over 1,500 metres. Modelling with the non-migratory population database alone might bias in terms of representation of the study cohort, but it would not involve bias in terms of modelling the environment.

The phase 2 study used the potential confounding variables, such as demographic, socio-economic variables and human behaviour data, to adjust the model in the analysis. The possible biases were controlled at the stage of study design and field operation. The results of multilevel Poisson regression analysis not only corrected the potential confounding effects, but also the spatial autocorrelation of malaria within the same households and same villages. Hence, the results of the phase 2 study were more reliable.

The negative correlation of altitude and malaria was noted centuries ago (Hirsch, 1883). Hirsch stated that extent and severity of malaria diseases diminish in proportion as we ascend above the sea level in America. The protection afforded by altitude was described by many traditional textbooks on malaria (MacDonald, 1957; Russell, 1963). Dutta and Dutt (1978) have reviewed the historical data on association of altitude and the risk of malaria in different places in the world. They concluded different species of mosquitoes have different ranges of altitude limits. In general, on and near the equator, because the sun's rays are more direct, a higher altitude is susceptible to malaria as compared with temperate regions.

The effect of altitude on the risk of malaria was noted by the first epidemiological study in Yunnan by Yao and his colleagues (Ling *et al.*, 1936). They observed that malaria infection was practically limited to small isolated plains bounded by high mountains, places where water collects and malaria mosquitoes breed, and the top of the mountains was usually free from the infection. The association between malaria and altitude in Yunnan was further assessed by many other researchers (Roberston, 1940; Williams, 1941; Wang *et al.*, 1957 & Zhang, 1990). All the studies showed a consistent result, the higher altitude and the lower endemicity of malaria. But most previous studies were descriptive studies with crude analyses, and subjected to the number of biases such as "ecological fallacy" and the probability of confounding effects of other variables as described in the phase 1 study in this work.

The effect of altitude on malaria may be mainly due to temperature (Molineaux, 1988), which critically affects both the breeding cycles of mosquitoes and the duration of the sporogonic cycle after they are infected with malaria parasites (MacDonald, 1957). Over a certain range of temperature, high temperature speeds up mosquito development, and decreases the time interval between blood meals, thereafter leading to more frequent, mosquito-man contact, and shortening the duration of sporogony (Molineaux, 1988). The altitude is highly correlated with temperature in Yunann as discussed in Chapter 2. Therefore, altitude can be used as the surrogate of temperature as an important spatial predictor of malaria transmission. Based on the results of present study, Malaria might be stratified into 4 strata in the study area, under 800 metres, highly endemic area, 801-1,200 metres middle endemic area and 1,201-1,500 metres low endemic area. The probability of malaria infection above 1,500 metres is very small as shown in Tables 5. 21 and 5.23.

6.2.2 The effect of paddy on the risk of malaria

The effect of paddy on the risk of *P. vivax* and *P. falciparum* was also analysed in the phase I and phase II studies. The results in the phase I showed a clear positive correlation between the risk of *P. vivax* and the amount of paddy in the administrative village; also for *P. falciparum*. The associations between malaria for both *P. vivax* and *P. falciparum* infection and paddy in the administrative village remained the same in the final multilevel model as shown in Table 4.12 and Table 4.13. Nevertheless, the assessment of the effect of paddy on the risk of malaria might be subject to a number of biases in the phase I study. The biases might be due to the nature of routine malaria data set (under reporting), outdated land use map, "ecological fallacy" in the analysis, and the unavailability of socio-economic and demographic variables for correction of the model etc as discussed in Chapter 4. Therefore firm conclusions can be drawn only after the biases are excluded in the model.

The phase II cohort study was designed and implemented to maximally eliminate biases we can identify. The results of the analyses were consistent with those in the phase 1 study. Paddy was positively correlated with *P. vivax* malaria infection in the Feng Chun Ling in the final ordinary and multilevel Poisson regression models. A similar trend was identified for the association between *P. falciparum* malaria and paddy in the study area, although it was not statistically significant, probably due to the relatively small number of falciparum malaria cases occurring in the study area. The consistent results of the two phases provide strong evidence that paddy would progressively increase the risk of *P. vivax* and *P. falciparum* malaria in the Red River basin area.

The association between paddy and malaria has been known for centuries. The term 'rice malaria' was coined in the 1930s to describe the association of rice cultivation in Europe with malaria (Service, 1989). It was believed that extension of rice fields had led to malaria outbreaks in America in the17th century and the subsequent decline of malaria was cited as linked to the decline of rice cultivation (Boyd, 1941; Kitron, 1987). Najera (1988), Service (1989) and Lacey & Lacey (1990) extensively reviewed the history of the association of rice and malaria and their mechanisms. In summary, rice fields are usually flooded for long periods, and they provide ideal breeding places for mosquitoes adapted to breeding in the rice field. Lindsay *et al.* (1995) found that individual children's exposure to bites of *An. gambiae* was associated with living adjacent to rice fields in a Gambian village, apart from other characteristics of their houses. Samarsinghe (1986) reported up to a 5-fold increase in malaria and 10-fold increase in *P. falciparum* cases from 1982 to 1985 in a newly irrigated rice cultivation area in Sri Lanka. In Indonesia, more than 80% of malaria cases were reported from the rice growing areas of Banjarnegara, Wonosobo and Purbolinggo where *An. aconitus* and *An. barbirostris* are the major vectors of malaria (Bang, 1987). In China, a strong correlation has been reported between the level of malaria morbidity and the amount of rice cultivation, where *An. sinensis* is a major malaria vector in the Yellow River basin and the Huihe River basin areas (Zhou, 1981; Liu *et al.*, 1995; Zhou, 1990). The recent decline of malaria in the Yellow River basin area is partly due to the decline of rice cultivation, apart from strong control measures (Shang & Hou, 1992).

There are, however, a few exceptions; for instance a statewise correlation analysis between malaria data (API) and rice cultivation from 1963 to 1983 showed no strong correlation between the two variables in India (Sharma *et al.*, 1994). In the Yagoua area in Cameroon despite rice cultivation, malaria has significantly decreased (Service, 1989). The improvements in living standards, greater use of mosquito nets and increased use of antimalarial drugs and insecticide were among the reasons for this. Nevertheless, most of the instances are ecological studies and only a very simple model with a single variable was in use for the analyses from which it would be very difficult to eliminate other biases due to the "ecological fallacy" and other confounding effects such as the use of mosquito nets, accessibility of antimalarial drugs and ongoing control programmes.

An. minimus and *An. sinensis* are the major vectors of malaria in the study area, Yunnan. The breeding sites of *An. sinensis* are basically in rice fields in Yunnan as in most areas in China (Zhou, 1981; Tang *et al.*, 1990). The breeding sites of *An. minimus* in Yunnan in general are in slow running water such as streams and in clear water bodies at the edge of forest in Yunnan (Zhou, 1991; Dong, 1993). Habitats also include in some places, terraced type paddy fields, seepage from springs, and small water reservoirs with clear water in the Southwest part of the province (Dong, 1991). Dong (1993) summarised the results of nine studies in Yunnan and concluded that around 20% of *An. minimus* are breeding in rice fields in the Southwest of Yunnan province, mostly in terrace paddy. The terrace paddy or terraced rice fields are a unique feature of rice growing in Yunnan. The terrace paddy is distributed around mountain side. It is irrigated by natural spring water, so the water in the terrace paddy is very clean and slow running from paddy fields at the top of mountain to the paddy in the lower part of mountain gradually. It makes an ideal place for *An. minimus* breeding in the province, particularly in the Red River basin areas (Dong, 1993). Therefore, the more the paddy around households would increase breeding sites and thereafter would increase the density of *An. minimus* and *An. sinensis*. More paddy progressively increased the risk of malaria and this is consistent with the theory of paddy providing suitable breeding places for *An. minimus* and *An. sinensis* in the study area. Subsequently, there is increased contact between the mosquitoes and man, and man would be at a higher risk of malaria both for *P. falciparum* and *P. vivax*.

There are, however, a few exceptions; for instance a statewise correlation analysis between malaria data (API) and rice cultivation from 1963 to 1983 showed no strong correlation between the two variables in India (Sharma *et al.*, 1994). In the Yagoua area in Cameroon despite rice cultivation, malaria has significantly decreased (Service, 1989). The improvements in living standards, greater use of mosquito nets and increased use of antimalarial drugs and insecticide were among the reasons for this. Nevertheless, most of the instances are ecological studies and only a very simple model with a single variable was in use for the analyses from which it would be very difficult to eliminate other biases due to the "ecological fallacy" and other confounding effects such as the use of mosquito nets, accessibility of antimalarial drugs and ongoing control programmes.

An. minimus and *An. sinensis* are the major vectors of malaria in the study area, Yunnan. The breeding sites of *An. sinensis* are basically in rice fields in Yunnan as in most areas in China (Zhou, 1981; Tang *et al.*, 1990). The breeding sites of *An. minimus* in Yunnan in general are in slow running water such as streams and in clear water bodies at the edge of forest in Yunnan (Zhou, 1991; Dong, 1993). Habitats also include in some places, terraced type paddy fields, seepage from springs, and small water reservoirs with clear water in the Southwest part of the province (Dong, 1991). Dong (1993) summarised the results of nine studies in Yunnan and concluded that around 20% of *An. minimus* are breeding in rice fields in the Southwest of Yunnan province, mostly in terrace paddy. The terrace paddy or terraced rice fields are a unique feature of rice growing in Yunnan. The terrace paddy is distributed around mountain side. It is irrigated by natural spring water, so the water in the terrace paddy is very clean and slow running from paddy fields at the top of mountain to the paddy in the lower part of mountain gradually. It makes an ideal place for *An. minimus* breeding in the province, particularly in the Red River basin areas (Dong, 1993). Therefore, the more the paddy around households would increase breeding sites and thereafter would increase the density of *An. minimus* and *An. sinensis*. More paddy progressively increased the risk of malaria and this is consistent with the theory of paddy providing suitable breeding places for *An. minimus* and *An. sinensis* in the study area. Subsequently, there is increased contact between the mosquitoes and man, and man would be at a higher risk of malaria both for *P. falciparum* and *P. vivax*.

6.2.3 The effect of forest on the risk of malaria

The association between forest and risk of malaria in the administrative villages was studied in the ecological study as shown in Chapter 4. Although it showed a trend of more forest and the higher risk of malaria, the association was not statistically significant. The interpretation of the results might be either that, because of the relative small number of administrative villages under study, there was not enough power to detect such an association in the study area. Alternatively the lack of association between two variables might be due to the quality of data, for instance, out of date land use map, malaria reporting variation among the administrative villages and unavailability of the confounding variables for correction of the model etc. "Ecological fallacy", might also result in the inconclusive results

In the prospective cohort study, a definite positive association between risk of *P. vivax* and forest in the study area emerged in all final models. There is also a general trend of positive correlation between risk of *P. falciparum* and forest, but the association is not statistically significant, probably due to few cases and inadequate power to get significant results. Nevertheless, we can conclude that the study revealed a strong association of risk of malaria and forest.

Although the effect of forest dramatically increased the risk of vivax malaria at the fourth quintile there was no further increasing at the fifth quintile as shown in Table 5.25 of Chapter 5. The results suggest that mosquitoes might need at least certain amount of forest (threshold) to maintain their density and malaria transmission. More forest doesn't progressively increase the risk of malaria, perhaps because there might be no linear correlation between the amount of forest and the length of forest edges, which provides the vital environment for *An. minimus* breeding in the study area (Dong, 1993). Forest edge may reach a maximum or plateau at forest well short of complete coverage.

Forest has been linked to malaria for centuries. The common theory of the origin of malaria is that the natural focus of co-evolution between man, mosquito and malaria parasites lies in the tropical forest and it spread afterwards to other places (Coatney, 1971; Bruce-Chwatt, 1988; Poolsuwan, 1995). Forest is particularly a problem for malaria in the Southeast Asia, as extensively reviewed in a book edited by Sharma and Kondrashin (1991). In Bangladesh, for instance, forest accounted for 16.4% of their total land, and forest-related malaria cases accounted for 80% of the total cases of malaria; 70% of the forest malaria was *P. falciparum* infection (Haq & Maheswary, 1991). In India, 54 million "tribes" of various ethnic origins live in the forest areas and account for 7% of total population contributed 30% of total malaria cases, of which 60% of them were *P. falciparum* infection in 1987 (Narasimham, 1991). In Myanmar, about 60% of all malaria cases are reported from forest and forest fringe areas while the remaining 32% cases are reported from non-endemic areas as a results of malaria diffusion or migration (Tin & Tun, 1991). Therefore, there is clear evidence of malaria associated with forest in the Southeast Asia region, particularly with *P. falciparum*. Nevertheless, all the studies above are based on routine reports. The "forest related" malaria is surrogate for a group of environmental, socio-economic and human behaviour variables which are aggregated in the forest, such as, deforestation, mining, migration, and poor health services etc, as discussed by Kondrashin *et al.*, (1991).

The difficulty of analysis of the effect of forest on the risk of malaria was due to the way of measuring of the amount of forest prior to availability of GIS and remotely sensed technology such as satellite imagery and GPS. Traditionally, one might crudely define a village or household within forest or outside of forest, or roughly the distance from the forest fringe (Kalra, 1991). It was not feasible with traditional malariological methods to accurately measure the amount of the forest around households, nor the exact distance between the household and forest fringe, especially when studying a large number of households spread over a very large area prior to availability of GIS and remote sensing.

The majority of *An. minimus* was breeding in slow running water of streams and in clear water bodies at the edge of forest in Yunnan (Zhou, 1991; Dong, 1993). A summary review by Dong (1993) showed that 55% to 85% of mosquito breeding was in the clear streams, mostly near the forest fringe. Therefore increasing the forest edge would lead to more breeding places and consequently increase malaria transmission.

6.2.4 Migration: one of main malaria determinants in the study areas

Around 14% of the study cohort had experience of temporary migration to high-risk areas at low altitude areas for mining, and agriculture during the 7-month study period. Most of the migrating population were adult males, and came from villages with an altitude over 1,200m, where malaria is not generally endemic. On the whole, the population with history of migration during the study period contributed 61.3% (400/652) of malaria cases in the study cohort. In the crude analysis, the relative risk of malaria of the migratory population was around 10 fold as compared with that of non-migratory people as shown in Table 5.11 although this comparison has a number of problems such as the comparability of the two groups by age and sex *etc.* The study provided strong evidence that temporary migration was the major malaria determinant in the study area, particularly for the population who lived in the high altitude areas.

Population migration has been long recognised as a major factor of increasing the risk of malaria. Ross (1911), for example, described the cases of malaria related to population movement. He defined them either as imported cases, when the infected person immigrates into the locality; or extraneous infected cases, where persons contract malaria from a locality, outside their usual habitation. In Africa, Walton (1947) found that the entomological data collected in Freetown, Sierra Leone, couldn't totally explain the relatively higher risk than expected among the young population, suggesting the majority of these infections where contracted the outside the local area. During the period of the global malaria eradication programme, WHO (1957) considered this a special problem that could be seriously impact the success of the eradication programme.

One of the early studies in Yunnan China indicated that malaria cases in Yunnan, especially the 'malignant type' due to *P. falciparum,* were clearly linked to travelling for trade to Myanmar (Faust, 1926). In 1919, a 'malaria' outbreak in Simao was due to infection being brought to the city by soldiers, who contracted it from the civilians on the border. Huge migration of population has acted as a driving force for malaria transmission in Yunnan in recent years (Zhu *et al.*, 1994). But most of the previous studies were of long term movement in Yunnan, as reviewed by Xu and Liu (1997). The present study focused on temporary migration among the local population within the study area. Higher risks among the people with a history of migration might have a number of causes. First movement from one set of environments to another might result in the exposure to vectors of malaria. Secondly, movement brings different groups of people into contact with one another and increases the possibility of transmission of diseases. Thirdly, physical or psychological stress might result from movement and therefore might make people more susceptible to the malaria. It seems clear that the first of these predominates in the study area.

6.2.5 Use of mosquito nets was protective

Mosquito nets have been long considered an effective method to protect against mosquito biting and reduce the risk of malaria (Lindsay & Gibson, 1988; Curtis, 1992). A surprising finding in the crude analysis of the phase 2 study, however, was that the users of mosquito nets were at a higher risk of malaria, particularly for *P. falciparum* malaria infection. Mosquito net users had 1.26 fold and 2.43 fold risk of *P. vivax* and *P. falciparum* indigenous malaria as compared with non-users of mosquito nets in the study area. More detailed analyses revealed that most mosquito net users resided in the lower altitude areas where mosquitoes were at the highest density and therefore malaria transmission was very intensive. The protective efficacies of mosquito nets against *P. vivax* and *P. falciparum* were 40% and 29%, respectively in the final multilevel Poisson regression models. The interpretation of the differences of the protective efficacy of mosquito nets on *P. vivax* and *P. falciparum* should be made with caution. It might be due to chance because the 95% confidence intervals of two point estimates of RRs (Table 5. 24 and Table 5.26) overlapped. The relatively small number of *P. falciparum* malaria cases make the point estimate of RR very unstable with a wide confidence interval. On the other hand, mosquito net users resided in the low altitude areas, and also the dominant parasite species was *P. falciparum* in the low altitude areas, so some temporarily migrants population might be infected with *P. falciparum* and then be misclassified as indigenous cases without nets in the high mountain area. Consequently, the misclassification could dilute the measured protective efficacy of mosquito net against *P. falciparum* infection.

Several studies have been investigated the effect of mosquito nets on malaria morbidity recently. Here, and in the field study, the effect of untreated mosquito net is being considered. In The Gambia, retrospective studies revealed mosquito nets had a strong protective effect on parasite parasitemia and spleen rates. After accounting for ethnic group and the places of residence, a significant inverse correlation between splenomegaly and the use of mosquito nets remained (Bradley *et al.*, 1986, Campbell *et al.*, 1987). But a further randomised control trial among The Gambia children indicated that there was no significant difference in the incidence of clinical attacks of malaria among the bed net users and non users although less clinical attacks were demonstrated among children using mosquito nets (Snow *et al.*, 1988). But the authors noted that a significant number of children left their nets for a period during the night. A larger study later in The Gambia revealed that untreated mosquito nets provide some individual protection against malaria, although not as efficiently as that provided by insecticide-treated mosquito nets (D'Alessandro *et al.*, 1995). That mosquito nets are protective is consistent with entomological studies which have found the mosquito nets deter biting by mosquitoes, resulting in reduced feeding success (Lines *et al.*, 1987) and a lower human blood index (Lindsay *et al.*, 1989; Burkot *et al.*, 1990). One interpretation of the protective effect of mosquito nets in the study cohort might be due to increasing infection in non-mosquito net user because of diversion of mosquitoes to these non users and not due to a reduction in malaria morbidity in the mosquito net users (Lines *et al.*, 1987). This possibility is small in the present study. Because only a small proportion of population in the study cohort were mosquito net users, and one of two vectors, *An. sinensis*, prefers animal blood in the study area, it is unlikely that the diversion of mosquitoes from mosquito net users to non-users caused these differences of the risk in malaria among mosquito net users and non-users. Another study in The Gambia (Thomson *et al.*, 1994) showed that mosquito net usage was strongly correlated with vector density and the highest malaria rates were found in villages where mosquito net use was relatively low. They concluded that in The Gambia malaria prevalence rates are reduced where nuisance biting by mosquitoes is sufficient to encourage the population to protect themselves with mosquito nets. The results of the present study are similar to those in The Gambia, although the overwhelming evidence is that insecticide-treated mosquito nets are protective and that mosquito nets are much less effective until treated with insecticide (Snow *et al.*, 1988; Lengeler, 1998), The present study, however, has shown that mosquito nets are protective and the protective effect was unlikely to be due to diversion of mosquitoes to non-users of mosquito nets.

6.3 Implications of the study findings for the control and surveillance of malaria

Yunnan is one of the most highly endemic areas for malaria in China, particularly in the Red River basin and its border areas. Traditional intensive surveillance systems for malaria are becoming unreliable. The terrain shows considerable variation and some of it is relatively inaccessible. The environmental or land use pattern, particularly in the lower altitude areas, keeps on changing due to overpopulation. Migration causes malaria outbreaks among migratory people themselves and the local population. The present study used GIS, integrated with satellite imagery data and malaria data to identify malaria risk factors related to environmental and human behaviour variables and to develop a predictive model in the hope of predicting malaria high risk areas and outbreaks, and guiding malaria control in the Red River basin area.

The findings of the study have shown that the spatial distribution of indigenous malaria was mainly determined by landscape and land use variables in the study areas, altitude, paddy and forest. The dominant determinant is altitude. The results of ordinary and multilevel modelling of Poisson regression analysis revealed that the population living in the areas with altitude above 800 metres had less than one fifth of the risk of malaria of those who resided in the area with altitude below 800 metres. We further tabulate malaria incidence rate against different altitude categories for the local none temporarily migratory population. As shown in Table 6.1, the areas below 800 metres had 8.5% of total population, but account for 31.7 % of total malaria cases among non-migrated population. The areas below 1,200 metres had only 35.2% of total population, but account for 62.3 % of total malaria cases among non-temporary migration population in the study area.

Table 6.1. Malaria distribution of non- temporary migrated population at different altitude categories

Altitude	Pop.	Pop. (%)	Cases	Incidence rate (%)	RR	Cumulated cases (n)	Cumulated cases (%)	Cumulated population (%)
<800	1,783	8.5	80	4.5	1.00	80	31.7	8.5
-1,200	5,579	26.7	79	1.4	0.31	157	62.3	35.2
-1,350	4,344	20.8	30	0.7	0.16	187	74.2	56.0
-1,500	6,140	29.4	47	0.8	0.18	236	93.7	85.4
>1,500	3,062	14.6	16	0.5	0.11	252	100	100

Paddy and forest also contributed to significant variation of malaria in the study area, particularly in the middle and high altitude areas in the study cohort. They provide vital environments for *An. sinensis* and *An. minimus* for breeding sites and increase the density of mosquitoes. The findings imply that the development and reuse of the land in the Red River basin area should be very careful under the current land pressure due to overpopulation in the middle and higher mountain zones. Future developers of land should be aware of the potential malaria and other vector borne disease risk arising from expansion of paddy fields, which will provide breeding sites for *An. sinensis* and *An. minimus*, particularly in the middle and lower altitude area. A major problem in Yunnan, China today is that health considerations usually rank well below agricultural priorities in re-used land for agriculture; there seems to be a desire to maximise food production and direct financial benefit almost at whatever cost, due to overpopulation in the province. Deforestation might increase the forest edges and increase the breeding sites for anopheline mosquitoes in the study area. The consequence of this kind of development of the land into paddy and deforestation would results in malaria outbreaks or epidemics in the Red River basin area as well as other border areas of Yunnan province, China. Yet if the health component is forgotten the full benefits of agricultural development cannot be fulfilled.

As discussed in last paragraph, environmental features, particularly land use patterns were continuing to change due to overpopulation in the Red River basin area today. A large amount of forest has been cut down in the area. Some land has been developed into paddy fields. It is impossible for the ground-based malaria worker to be aware of where the forest have been cut down and where the lands have been developed into paddy fields in a vast area like the whole Red River basin area. Satellite remote sensing provides continuing measurements of environmental and ecological variables on a relatively short temporal cycle making it particularly valuable for monitoring such changes in the area. They would quickly help malaria epidemiologists and malaria control authorities to be aware of where new paddy has been developed and forest has been cut down in the area. By the integration of land use change from a satellite imagery data and altitude data in a GIS with the spatial predictive model developed in this study, they can identify malaria risk and likelihood of malaria outbreaks in the changed environment areas to some degree. Then necessary malaria control resources should allocated in time.

The finding of the studies have also shown that human behaviour plays a very important role in malaria distribution in the study area. People with a history of temporary migration were found to have a much higher risk of malaria in the study area, particularly for *P. falciparum* malaria infection. 14% of the study cohort with a history of temporary migration accounted for 61.3% of total malaria cases in the study cohort. People with a history of temporary migration had 10-fold higher risk of malaria as compared with that of the non-migratory population during the study period. A further danger of the temporary migration is that those people could bring malaria parasites back to their own villages leading to malaria transmission and even outbreaks if the landscape and land use conditions favour malaria transmission (Xu & Liu, 1997; Zhang *et al.,* 1992). Therefore, specific malaria control strategies should be studied to prevent this highest malaria risk but a most productive group in the Red River basin area. Temporarily migrating populations, however, are very mobile yet with little or no immunity to malaria. The places they work are at the low altitude areas where malaria is extensively transmitted. Their residence places are usually temporary huts, not effectively protected by spraying an insecticide such as DDT, which is currently practised malaria control policy in the Red River basin area, Yunnan. Chemoprophylaxis and other control measures need to be explored for the highest risk groups in the Red River basin area although it would have the danger of selection pressure for the emergence of resistance (Gilles & Warrell, 1993).

Mosquito nets provided significant protection against indigenous malaria in the study cohort. The protective efficacy for *P. vivax* and *P. falciparum* were 40% and 29%, respectively. If all the population used mosquito nets, it would decrease malaria risk in the Red River basin area. Nevertheless, only less than one fifth of the total population in the study area own mosquito nets, although 57% of population in the area with altitude below 800 metre used mosquito nets. The first priority of malaria control strategy in the immediate future is to encourage local residents and 'downhill' migrants to use mosquito nets in the Red River basin area and to ensure these are regularly treated with insecticides.

The complexity of the modelling, and especially the need to collect primary data in the field to compensate for the limitation of routine surveillance, have limited the degree of other exploration of the model's use for control proposes, but multilevel modelling proved a useful tool for handling the variation in reporting by area, in seeking environmental determinants.

The finding that migration was the predominant single determinant of malaria risk shows a clear limitation on the value of spatial modelling as a guide to control programmes: the map can show areas of high transmission risk but will be much less good to indicate places with concentrations of patients. Moreover, the areas with a very high relative risk for indigenous acquired malaria have a limited population and therefore do not account for the preponderant case loads. Therefore, malaria control strategy making should be based on information on the high transmission areas identified by GIS and remote sensing modelling and on high risk population groups from epidemiological study. In present cases, we might use GIS to map high malaria transmission zones based on information on altitude, paddy and forest, but specific attention should be paid to the temporary migrant group in future malaria control programmes.

6.4 Future research

The present study grew to a large undertaking, due to the need for primary incidence data of high quality. However it still raised many further questions beyond its current scope. Such future research could include the following:

1. ***The model validation and malaria mapping and prediction in the Red River basin area.*** A further study should be carried out in other places of the Red River basin area, so that the model developed in the phase II study could be validated. After the validation of the model, a predictive malaria transmission map should be produced in the Red River basin area based on the spatial predictive model. Nevertheless, such a validating study needs re-running with a much large population of 'gold standard', upgraded routine surveillance in a relatively large area by local malaria control authorities to obtain more accurate malaria data that would be more cost-effective in terms of testing the model. Similarly, a spatial predictive model of malaria should be developed by using GIS and remotely sensed data in the border areas of Yunnan, where malaria is highly endemic, and the areas are different from the Red River basin in terms of landscape and ecology. The models being developed for the Red River basin and border areas in Yunnan might help other countries, particularly in Southeast Asia to understand malaria transmission and its spatial distribution patterns in their countries, such as Thailand, P.R. Lao, Vietnam, Cambodia and Myanmar, which have similar landscape and ecological environmental situations to the Red River basin and border areas of Yunnan.

2. *Malaria and its vector mapping in Yunnan.* Based on the results of present study, the previous literature as reviewed in Chapter 2, the preliminary GIS mapping on malaria in Yunnan (Hu *et al.*, 1998; Kidson *et al.*, 1999), malaria is strongly correlated with altitude and vegetation (forest and rice) in Yunnan. Malaria and its vector spatial distribution should be able to be mapped based on the two variables in the province. Low resolution satellite imagery data, such as AVHRR-NOAA and Meteosat-HRR, can be used to map the vegetation and meteorological conditions in the province. By integration of vegetation and meteorological data derived from AVHRR-NOAA or Meteosat-HRR, meteorological data on the ground, digital elevation model (DEM) of the province and malaria data in Yunnan into a GIS, we can model the malaria and its vector spatial distribution. Subsequently a malaria risk map should be developed in the province. Nevertheless, meteorological satellite data will only help map malaria and its vector with lower resolution in the province, but the local variation of malaria will need the higher resolution satellite data such as Landsat TM data and SPOT imagery data as described in the present study. Mapping malaria and its vector in Yunnan as a whole should combine the theoretical (experimental approach) model of climate suitability of malaria transmission and an empirical model of the disease (regression approach).

3. *Malaria drug resistance mapping in Yunnan.* Anti-malarial drug resistance is a serious problem of malaria control in Yunnan. But only very limited data are available to make up the picture of anti-malaria drug resistance as shown in Figure 2.4 of Chapter 2. Using a GIS to integrate the drug resistance data and other data sources might help to map and predict the spatial distribution of drug resistance malaria in the Province once a more substantial body of data is gathered in the field. The first step of the mapping should develop a climate suitability model and empirical model to map *P. falciparum* malaria distribution in Yunnan. And the second step might use a regression approach or interpolation approach such as kriging to predicted the detail spatial distribution of drug resistance to *P. falciparum* in the province and even the whole Southwest of China.

4. *Researching on improving use of mosquito nets.* The methodology to encourage insecticide treated mosquito nets should be developed in the Red River basin area. In order to achieve the objectives, the socio-economic behaviour of the population related to the use of mosquito nets will need to be studied by a social scientist.

5. ***The assessment of the efficacy of insecticide treated mosquito net against malaria in Yunnan, China.*** Although the use of mosquito nets was found to be effective against malaria in the present study, the protective efficacy of mosquito nets might be enhanced by treatment with insecticide (Snow *et al.*, 1988; Lindsay *et al.*, 1989; Burkot *et al.*, 1990; D'Alessandro *et al.*, 1995). Only very limited data were available on the efficacy of insecticide treated mosquito nets against malaria and its vectors in Yunnan (Zhang & Yang, 1996) although considerable studies on mosquito nets treated with insecticides were carried out in other parts of China (Luo *et al.*, 1994a; Luo *et al.*, 1994b; Luo *et al.*, 1996; Guo *et al.*, 1996). The efficacy of mosquito nets impregnated with pyrethroid insecticide should be studied as a tool to control malaria in the Red River basin area. A village- based randomised control trial should be designed and implemented to assess the efficacy of mosquito nets impregnated with pyrethroid insecticide against malaria, *An. sinensis* and *An. minimus* in the Red River basin area and also other areas of Yunnan where people have different customs and behaviours.

6. ***Malaria control strategy study on temporarily migrating populations.*** An effective surveillance and preventive strategy for malaria in the migratory population should be sought urgently. One of the strategies for malaria control for temporary migration population to high-risk area is chemoprophylaxis. Little study has been done on the protective efficacy of chemoprophylaxis in Yunnan, or on the use of mosquito nets by these migrants. The potential effect of temporary migration on the risk of malaria in local villages from which the migrants come should also be a priority for the future research.

6.5 Summary and conclusion

The malaria situation has been deteriorating in the Red River basin, Yunnan in recent years as a result of environmental changes, deforestation, reused land, and increasing migration in the region. The present study used GIS and remotely sensed satellite data to identify malaria risks related to landscape and environmental factors, in the hope of finding useful environmental indicators to predict in which areas resources should be concentrated to have maximum benefit in the Red River basin area.

The study was carried out in Yuanyang County of the Red River basin area, Yunnan, China. The study was divided into two phases. The phase I study was a retrospective study. Malaria and population data in unit of administrative village were collected in 131 administrative villages of the county. Multilevel logistic regression models were used to model the malaria data and landscape and environmental variables derived from GIS over a series of maps. The results of the analyses revealed that malaria spatial distribution of malaria was determined by their altitude and land use pattern in the administrative villages. Malaria was negatively correlated with mean altitude of administrative villages, but paddy in the administrative villages would increase the risk of malaria in the villages.

Phase II study is a prospective study. The study was carried out in the Feng Chun Ling Township of Yuanyang County. The population of 24,280 in 5,007 households were included in the study. All household locations (identified by GPS), updated land use map (derived from a SPOT 4 image) and malaria databases were integrated into a GIS. GIS spatial analysis function was used to quantify altitude and land use variables in the study cohort.

The results of multilevel Poisson regression analysis revealed that the mosquito nets were protective for the malaria. The protective efficacy for *P. vivax* was 40% and *P. falciparum* 29%, respectively. The people with history of temporary migration during study had around 10-fold risk for malaria as compared with those of population without any history of migration in a crude analysis. The analyses also indicated that malaria was negatively correlated with altitude for both *P. vivax* and *P. falciparum.* But more paddy and forest would increase the risk of malaria, although the effect of forest on malaria was not linearly correlated. Once the amount of forest reached a certain threshold, no further increasing of the risk of malaria occurred. Presumably, the amount of forest edge for mosquito breeding was not linearly correlated with the amount of forest. Deforestation would increase the forest edge in the dense forest area, consequently would tend to increase mosquito density.

In conclusion, malaria transmission in the study area was mainly determined by the environmental variables particularly altitude, paddy and forest in the Red River basin area. But human behaviours such as the use of mosquito nets and temporary migration also play a very important role in the spatial pattern of malaria in the study area. Malaria control and surveillance should focus on the lower altitude area and on the mobile population in the Red River basin area. The overall temporarily migrated population plus population living below 1,200 metres accounted for 44.2% (10734/24280) of total population, but accounted of 85.7% (559/652) of total malaria cases during the study period, while migrant plus those living under 800 metres account for 73.6% (480/652) of cases in only 21.2% (5155/24280) of the population as shown in Table 6.2. Therefore, malaria control and surveillance should focus on population living in the lower altitude areas and temporarily migrating population. Future development of land should be aware of the potential malaria and other vector borne disease risk arising from expansion of paddy field and deforestation, which will provide breeding sites for mosquitoes, particularly in the middle and lower altitude areas. Subsequently, they might result in malaria and other vector borne disease outbreak or epidemics. Mosquito nets should be encouraged for use in the Red River basin area. Chemoprophylaxis might also be considered. Certainly control measures need to be explored for the temporarily migrating population in the Red River basin area.

Table 6.2. Cumulative malaria distribution based on altitude and temporary migration variables in the study cohort

	Pop.	Pop .(%)	Cases	Incidence per 1000	RR	Cumulated cases (n)	Cumulated cases (%)	Cumulated pop. (%)
Migrating	**3,372**	**13.9**	**400**	**118.6**	**17.2**	**400**	**61.3**	**13.9**
Indigenous 800 m	**1,783**	**7.3**	**80**	**44.9**	**6.5**	**480**	**73.6**	**21.2**
Indigenous 800 –1,200m	5,579	23.0	79	14.2	2.1	559	85.7	44.2
Remainder	13,546	55.8	93	6.9	1.00	(93)	(14.3)	(55.8)
Total	24,280	100	652	26.9	0.11	252	100	100

References

Advisory Committee on Malaria, MoH. (1993). Malaria situation in the People's Republic of China in 1992. *Chinese Journal of Parasitology and Parasitic Diseases,* **11**,161-164.

Advisory Committee on Malaria, MoH. (1994). Malaria situation in the People's Republic of China in 1993. *Chinese Journal of Parasitology and Parasitic Diseases,* **12**,161-164.

Advisory Committee on Malaria, MoH. (1995). Malaria situation in the People's Republic of China in 1994. *Chinese Journal of Parasitology and Parasitic Diseases,* **13**,161-164.

Anyamba, A. & Eastman, J. R. (1996). Interannual variability of NDVI over Africa and its relation to EL Nino/ Southern Oscillation. *International Journal of Remote Sensing,* **17,** 2533-2548.

Armitage, P. & Berry, G. (1987). *Statistical Methods in Medical Research.* 2nd, Oxford: Blackwell Scientific Publications.

Aronoff, S. (1989). *Geographical Information Systems: A Management Perspective.* Ottawa, Canada: WDL Publications.

Bang, Y. H. (1987). *Vector borne diseases problems associated with rice cultivation acreage and their control of Southeast Asia.* FAO publication, AGL/MISC/12/87: 116-125.

Barnes, C. M. & Cibula, W. G. (1979). Some implication of remote sensing technology in insect control programs including mosquitoes. *Mosquito News,* **39,** 271-282.

Barradell, L. B. & Fitton, A. (1995). Artesunate. A review of its pharmacology and therapeutic efficacy in the treatment of malaria. *Drugs,* **50,** 714-741.

Barrera, R., Grillet, M. E., Rangel, Y., Berti, J. & Ache, A. (1999). Temporal and spatial patterns of malaria reinfection in Northeastern Venezuela. *Amercian Journal of Tropical Medicine and Hygiene,* **61,** 784-790.

Beck, L. R., Washino, R. K., Hibbard, K. A. & Salute, J. S. (1992). Estimating high mosquito-producing rice field using spectral and spatial data. *International Journal of Remote Sensing,* **13,** 2813-2826.

Beck, L. R., Rodriguez, M. H., Dister, S. W., Rodriguez, A. D., Rejmankova, E., Ulloa, A., Meza, R. A., Roberts, D. A., Paris, J. F., Spanner, M. A., Washino, R. K., Hacker, C. & Legters, L. J. (1994). Remote sensing as a landscape epidemiological tool to identify villages at high risk from malaria transmission. *American Journal of Tropical Medicine and Hygiene,* **51,** 271 - 280.

Beck, L. R., Rodriguez, M. H., Dister, S. W., Rodriguez, A. D., Washino, R. K., Roberts, D. A. & Spanner, M. A. (1997). Assessment of a remote sensing-based model for predicting malaria transmission risk in villages of Chiaps, Mexico. *Amercian Journal of Tropical Medicine and Hygiene,* **56,** 99-106.

Binka, F. N. & Smith, T. (1998). Impact of spatial distribution of permethrin-impregnated bed nets on child mortality in rural northern Ghana. *Amercian Journal of Tropical Medicine and Hygiene,* **59,** 80-85.

Bonham-Carter, G. F. (1994). *Geographic information systems for geoscientists: modelling with GIS.* New York: Pergamon.

Boundin, C., Robert, V., Carnevale, P. & Ambroise-Thoms, P. (1992). Epidemiology of *Plsamodium* in rice field and a savanna area in Burkina Faso. Comparative study on the acquired immunoprotection in native populations. *Acta Tropica,* **51,** 103-111.

Bouma, M. J. & van der Kaay, H. J. (1994). Epidemic malaria in Indian and the EL Nino Southern Oscillation. *Lancet,* **344,** 1638-1639.

Bouma, M. J. & van der Kaay, H. J. (1996). The EL Nino Southern Oscillation and historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics? *Tropical Medicine and International Health,* **1,** 86-96.

Bouma, M. J., Poveda, G., Rojas, W., Chavasse, W., Quinones, M., Cox, J. & Patz, J. A. (1997). Predicting high-risk year for malaria in Colombia using parameters of EL Nino Southern Oscillation. *Tropical Medicine and International Health,* **2,** 1122-1127.

Boyd, M. E. A (1941). A historical sketch of the prevalence of malaria in North America. *American Journal of Tropical Medicine.* **21,** 223-244.

Bradley, A. K., Greenwood, B. M., Greenwood, A. M., Marsh, K., Byass, P., Tulloch, S. & Hayes, R. (1986). Bed-nets (mosquito-nets) and morbidity from malaria. *Lancet,* **2,** 204-207.

Bradley, D. J. (1993). Human tropical diseases in a changing environment. In: *Environmental Change and Human Health,* Ciba Foundation syposium No. 175, pp.146. Chichester, UK. Wiley & Sons Ltd.

Bradley, D. J. (1994). Watson, Swellengrebel and species sanitation: environmental and ecological aspects. *Parassitologia,* **36,** 137-147.

Bradley, D. J. (1999). The last and the next hundred years of malariology. *Parassitologia,* **41,** 11-18.

Briggs, D. J. (1992). Mapping environmental exposure. In: *Geographical and environmental epidemiology, methods for small-area studies,* Elliot, P., Cuzick, J., English, D. & Stern, R. (editor). Oxford, pp. 158-176. Oxford University Press.

Bruce-Chwatt, L. J. (1991). *Essential malariology,* London: Edward Arnold

Bruce-Chwatt, L. J. (1988). History of malaria from pre-history to eradication. In *Malaria: principles and practice of malariology,* Wernsdorfer, W. H. & McGregor, I. (editor) Vol. 1. Pp. 1-59. Churchill, Livingstone.

Burkitt, D. (1962). A children's cancer dependent on climate factors. *Cancer,* **28,** 3 - 13.

Burkot, T. R., Garner, P., Paru, R., Dagoro, H., Barnes, A., McDougall, S., Wirtz, R. A., Campbell, G. & Spark, R. (1990). Effects of untreated bed nets on the transmission of *Plasmodium falciparum*, *P. vivax* and *Wuchereria bancrofti* in Papua New Guinea. . *Transaction of the Royal Society of Tropical Medicine and Hygiene*, **84**, 773-779.

Burrough, P. A. (1986). *Principles of Geographical Information Systems for Land Resources assessment*. Oxford: Oxford University Press.

Burrough, P. A. & McDonnell, P. A. (1998). *Principles of Geographical Information Systems*. Oxford: Oxford University Press.

Campbell, H., Byass, P. & Greenwood, B. M. (1987). Bed-nets and malaria suppression. *Lancet*. **1**, 859-860.

Carrat, F. & Valleron, A. J. (1992). Epidemiological mapping using the "kriging" method: application to an influenza-like illness epidemic in France. *American Journal of Epidemiology*, **135**, 1293-1300.

Chadee, D. D. & Kitron, U. (1999). Spatial and temporal patterns of imported malaria cases and local transmission in Trinidad. *Amercian Journal of Tropical Medicine and Hygiene*, **61**, 513-517.

Che, L. G., Chen, W. C., Yang, H. L. & Huang, K. G. (1986). The distribution of chloroquine resistant *P. falciparum* strains in Yunnan province. *Chinese Journal of Epidemiology*, **7**, 88-91.

Che, L. G., Li, X. L., Yang, Y. M., Li, C. F. & Zhang, Y. L. (1992). A survey on an outbreak of malaria at the Jiabutuo rubber plantation area, Yunnan. *Chinese Journal of Parasitic Disease Control*, **5**, 81-83.

Chen, C., Tang, L. H. & Chun, J. (1992). Studies on a new antimalarial compound: pyronaridine. *Transaction of the Royal Society of Tropical Medicine and Hygiene*, **86**, 7 - 10.

Chow, C. Y. & Balfour, M. C. (1949). The natural infection and seasonal prevalence of Anopheles mosquito in Chefeng and vicinity, Yunnan-Myanmar Border. *Chinese Medical Journal*, **67**, 405-413.

Clayton, D. & Hills, M. (1993). *Statistical Models in epidemiology*. Oxford: Oxford University Press.

Cliff, A. D. & Haggett, P. (1988). *Atlas of disease distributions: analytic approaches to epidemiological data*. Oxford: Blackwell Reference.

Cliff, A. D. & Haggett, P. (1996). The impact of GIS on epidemiological mapping and modeling. In. *Spatial analysis: modeling in a GIS environment*, Longley P & Batty M (editor). Cambridge, UK, pp. 321-344. GeoInformation International.

Cline, B. L. (1970). New eyes for epidemiologists: aerial photography and other remote sensing techniques. *American Journal of Epidemiology*, **92**, 85-89.

Coatney, G. R., Cooper, C. W. & Yang, M. D. (1950). Studies in human malaria XXX: A summary of 204 sporozoite-induced infections with chosen strain of *Plasmodium vivax*. *Journal of the National Malaria Society*, **9**, 381-396.

Coatney, G. R., Collin, W. E., Warren, W. M. & Contacos, P. G.(1971). *The primates malaria*. Bethgesda, USA

Coetzee M., Craig, M. H. & le Sueur D. (2000). Mapping the distribution of members of the *Anopheles gambiae* complex in Africa and adjacent islands. *Parasitology Today*, **16**, 74-77.

Cousens, S. N., Linsell, L., Smith, P.G., Chandrakumar, M., Wilesmith, J. W., Knight, R. S., Zeidler, M., Stewart, G. & Will, R. G. (1999). Geographical distribution of variant CJD in the UK. *Lancet*, **353**, 18-21.

Craig, M. H., Snow, R. W. & le Sueur, D. (1999). A climate-based distribution model of malaria transmission in Africa. *Parasitology Today*, **15**, 105-111.

Cross, E. R., Newcomb, W. W. & Tucker, C. J. (1996). Use of weather data and remote sensing to predict the geographical and seasonal distribution of *Phlebotomus papatasi* in Southeast Asia. *Amercian Journal of Tropical Medicine and Hygiene*, **54**, 530-536.

Cumming, G. S. (1999). Host distribution do not limit the species ranges of most African ticks (*Acari: Ixodida*). *Bulletin of Entomological Research*, **89**, 303-327.

Curtis, C. F. (1992). Prospects for the application of pyrethroid-impregnated nets or curtains to large-scale malaria control. In: *Appropriate Technology in Vector Control*, Curtis, C. F. (editor). Boca Raton, Florida: CRC Press, P. 38-43

Dale, P. E. R. & Morris, C. D. (1996). *Culex annulirostris* breeding sites in urban areas – using remote sensing and digital image analysis to develop a rapid predictor of potential breeding areas. *Journal of the American Mosquito Control Association*, **12**, 316-320

D'Alessandro, U., Olaleye, B. O., McGuire, W., Thomson, M. C., Langerock, P., Bennett, S. & Greenwood, B. M. (1995). A comparison of the efficacy of insecticide-treated and untreated bed nets in preventing malaria in Gambian children. *Transaction of the Royal Society Tropical Medicine and Hygiene*, **89**, 596-598

DeMers, M. N (1997). *Fundamentals of Geographic Information Systems*, New York: John Wiley & Son.

Diamond, I., Clements, S., Stone, N. & Ingham, R. (1999). Spatial variation in teenage conceptions in south and west England. *Journal of the Royal Statistical Society, Series A*. **162**, 273-289.

Dister, S. W., Fish, D., Bros, S. M., Frank, D. H. & Wood, B. L. (1997). Landscape characterization of peridomestic risk for Lyme disease using satellite imagery. *American Journal of Tropical Medicine & Hygiene*, **57**, 687-92.

Dong, X. S. (1993). *Mosquito Fauna, of Yunnan, China*, Vol 1, Kunming: Yunnan Scientific Public Press.

Dunn, C. E., Woodhouse, J., Bhopal, R. S. & Acquilla, S. D. (1995). Asthma and factory emissions in northern England: addressing public concern by combining geographical and epidemiological methods. *Journal of Epidemiology and Community Health,* **49,** 395-400.

Dutta, H. M. & Dutt, A. K. (1978). Malaria ecology: a global perspective. *Social Science and Medicine,* **12,** 69-84

Ecob, R. (1996). A multilevel modelling approach to examining the effects of area of residence on health and functioning. *Journal of the Royal Statistical Society. Part A,*.**159,** 61-75.

Elliot, P., Cuzick, J., English, D. & Stern, R. (editor). (1992). *Geographical and environmental epidemiology, methods for small-area studies,* Oxford: Oxford University Press.

English, D. (1992). Geographical epidemiology and ecological studies. In: *Geographical and environmental epidemiology, methods for small-area studies,* Elliot, P., Cuzick, J., English, D. & Stern, R. (editor). Oxford: Oxford University Press, pp. 3-13.

Estrada-Pena, A. (1998). Geostatistics and remote sensing as predictive tools of tick distribution: a cokriging system to estimate *Ixodes scapularis (Acari: Ixodidae)* habitat suitability in the United States and Canada from advanced very high resolution radiometer satellite imagery. *Journal of Medical Entomology.* **35,** 989-95.

Faust, E. C. (1926). An inquiry into the prevalence of malaria in China. *Chinese Medical Journal,* **40,** 937 - 956.

Fonaroff, L. S. (1982). Malaria ecology – the Trinidadian aspect. *Social Science and Medicine* **11,** 1-11.

Garnham, P. C. C. (1980). Malaria in its various vertebrate host: In: *Malaria,* Vol.1, Kreier J. P. (editor). Academic Press, p95-144.

Genton, B., Hii, J., al-Yaman, F., Paru, R., Beck, H. P., Ginny, M., Dagoro, H., Lewis, D. & Alpers, M. P. (1994). The use of untreated bednets and malaria infection, morbidity and immunity. *Annals of Tropical Medicine & Parasitology,* **88,** 263-70.

Ghebreyesus, T. A., Haile, M., Witten, K. H., Getachew, A., Yohannes, A. M., Yohannes, M., Teklehaimanot, H. D., Lindsay, S. W. & Byass, P. (1999). Incidence of malaria among children living near dams in northern Ethiopia: community based incidence survey. *British Medical Journal.* **319,** 663-666.

Ghebreyesus, T. A., Witten, K. H., Getachew, A., O'Neill, K., Bosman, A. & Teklehaimanot, A. (1999). Community-based malaria control in Tigray, Northern Ethiopia. *Parassitologia,* **41,** 367-371.

Gill, G. A. (1923). The prediction of malaria epidemics with special reference to actual forecast in 1921. *Indian Journal of Medical Research,* **10,** 1136-1143.

Gilles, H. M. & Warrell, D. A. (editor). (1993). *Bruce Chwatt's Essential Malariology.* London: Edward Arnold.

Glass, G. E., Amerasinghe, F. P., Morgan, J. M. & Scott, T. W. (1994). Predicting *Ixodes scapularis* abundance on white-tailed deer using geographical information systems. *Amercian Journal of Tropical Medicine and Hygiene,* **51,** 538 - 544.

Glass, G. E., Schwartz, B. S., Morgan, J. M., Johnson, D. T., Peter, M. A., Noy, M. & Israel, E. (1995). The environmental risk factors for lyme disease identified with geographical information systems. *American Journal of Public Health,* **85,** 944 - 948.

Gleiser, R. M., Gorla, D. E. & Almeida, F. F. L. (1997). Monitoring the abundance of *Aedes (Ochlerotatus) albifasciatus* (Macquart 1838) *(Diptera: Culicidae)* to the south of Mar Chiquita lake, central Argentina, with the aid of remote sensing. *Annals of Tropical Medicine and Parasitology,* **91,** 917-926.

Glynn, J. R. (1994). Infecting dose and severity of malaria: a literature review of induced malaria. *Journal of Tropical Medicine and Hygiene.* **97,** 300-316.

Glynn, J. R. & Bradley, D. J. (1995). Inoculum size and severity of malaria induced with *Plasmodium ovale. Acta Tropica,* **59,** 65-70.

Glynn, J. R. & Bradley, D. J. (1995). Inoculum size, incubation period and severity of malaria. Analysis of data from malaria therapy. *Parasitology,***110,**7-19.

Grilli, M. P. & Gorla, D. E. (1997). The spatio-temporal pattern of *Delphacodes kuschi* (Homoptera: Delphacidae) abundance in central Argentina. *Bulletin of Entomological Research,* **87,** 45-53.

Goldstein, H. (1995). *Multilevel Statistical Models.* 2nd, London: Edward Arnold.

Goldstein, H. & Rasbash, J. (1996). Improved approximation for multilevel models with binary response. *Journal of the Royal Statistical Society, A.* **159,** 505-513.

Goldstein, H., Rasbash, J., Plewis, I., Draper, D., Browne, W., Yang, M., Woodhouse, G. Healy, M. (1998). *A user's guide to MlwiN.* Multilevel Models Project, Institute of Education, University of London.

Guo, H. P., Zhu, Y., Dai Q., Yang, J. H., Jiang, K. X. & Zhong, M. (1998). The *in vivo* sensitivity of *Plasmodium falciparum* to chloroquine in the Red River basin, Yunnan, China. *Southeast Asian Journal of Tropical Medicine and Public Health,* **29.** 692-695.

Ge, Y. F., Cai, R. X. & Luo, D. P. (1996). Pyrethroid insecticide treated bednets for malaria control in the people's Republic of China. *Southeast Asian Journal of Tropical Medicine and Public Health,* **27,** 841-843

Harinasuta, H. & Bunnag, D. (1988). The clinical features of malaria. in *Malaria: Principles and Practice* (eds Wernsdorfer WH & McGregor I), Churchill Livingstone, London, pp 709-734.

Haq, M. I. & Maheswary N. P. (1991). Forest related malaria in Bangladesh. In: *Forest malaria in Southesat Asia,* Sharma, V. P. & Kondrashin, (editor). New Delhi.

Hay, S. I., Tucker, C. J., Rogers, D. J. & Packer, M. J. (1996). Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Annals of Tropical Medicine and Parasitology,* **90,** 1-19.

Hay, S.I., Snow, R.W. and D.J. Rogers. (1998a). From predicting mosquito habitat to malaria seasons using remotely sensed data: practice, problems and perspectives. *Parsitiology Today* **14,** 306-313.

Hay, S.I., Snow, R.W. and D.J. Rogers. (1998b). Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. *Transaction of the Royal Society Tropical Medicine and Hygiene.* **92,**12-20.

Hay, S.I., Packer, M.J., & Rogers. D. J. (1997). The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *International Journal of Remote Sensing,* **18,** 2899-2930.

Hay, S.I. & Lennon, J.J. (1999). Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. *Tropical Medicine and International Health,* **4,** 58-71.

Hayes, R.O., Maxwell, E.L., Mitchell, C.J. & Woodzick, T.L. (1985). Detection, identification, and classification of mosquito larval habitats using remote sensing scanners in earth-orbiting satellites. *Bulletin of the World Health Organization,* **63,** 361-374.

Hightower, A.W., Ombok,-M., Otieno, R., Odhiambo. R., Oloo, A. J., Lal, A. A., Nahlen, B. L. & Hawley, W. A. (1998).A geographic information system applied to a malaria field study in western Kenya. *American Journal of Tropical Medicine & Hygiene,* **58,** 266-272.

Hirsch, A. (1883). *Handbook of geographical and historical pathology,* Vol. London: The New Sydenham Society.

Hu, H., Singhasivanon, P., Salazar, N. P., Thimasarn, K., Li, X., Wu, Y., Yang, H., Zhu, D., Supavej, S. & Looarecsuwan, S. (1998) Factors influencing malaria endemicity in Yunnan Province, PR China (analysis of spatial pattern by GIS). Geographical Information Systems. *Southeast Asian Journal of Tropical Medicine and Public Health,* **29,** 191-200.

Huo Rong Huo. (1984), Malaria stratification in Yunnan. *Bulletin of the Yunnan Institute of Malaria Prevention and Treatment.* Vol. 1984, 1 –3.

Howe, G. M. (1986). Disease mapping. In: *Medical geography, progress and prospect,* Pacione M (editor). London: Croom Helm., pp. 35-63

Hsu, S. C., Sweet, W. C. & Balfour, M. C. (1949). Malaria in Chinese and Shan, Chefang, Yunnan-Myanmar border, *Chinese Medical Journal,* **67,** 1-12.

Hugh-Jones, M. (1989). Applications of Remote sensing to the identification of the habitats of parasites and disease vectors. *Parasitology Today,* **8,** 224-251.

Jones, K. & Ducan, C. (1996). People and place: the multilevel model as a general framework for quantitative analysis of geographical data, in *Spatial Analysis: Modelling in a GIS Environment*, Longley P & Batty M (editor), Cambridge: Geoinformation international.

Kalra, N. L. (1991). Forest malaria vectors in India: ecological characteristics and epidemiological implication. In: *Forest malaria in Southesat Asia,* Sharma, V. P. & Kondrashin, A. V (editor). New Delhi.

Karbwang, J. & Harinasuta, T. (1993). International border malaria control. *Southeast Asian Journal of Tropical Medicine and Public Health,* **25,** 615-616.

Kidson, C., Singhasivanon, P. & Supavej, S. (1999). Mekong malaria, malaria, multi-drug resistance and economic development in the greater Mekong subregion of Southeast Asia, incorporating geographical information systems databases. *Southeast Asian Journal of Tropical Medicine and Public Health,* **30** (suppl. 4), 1-101.

Kitanidis, P. K. (1997). *Introduction to geostatistics, application in hydrogeology.* Cambridge: Cambridge University Press.

Kitron, U. (1987). Malaria, agriculture, and development: lessons from past campaigns. *International Journal of Health Service.* **17,** 295-326.

Kitron, U., Jones, C. J., Bouseman, J. K., Helson, J. A. & Baumgartner. (1992). Spatial analysis of the distribution of *Ixodex dammini* (*Acari: Ixodidae*) on white-tailed deer in Ogle county, Illinois. *Journal of Medical Entomology,* **29,** 259 - 266.

Kitron, U., Pener, H., Costin, C., Orshan, L., Greenberg, Z. & Shalom, U. (1994). Geographic information system in malaria surveillance: mosquito breeding and imported cases in Israel, 1992. *American Journal of Tropical Medicine and Hygiene,* **50,** 550-556.

Kitron, U. & Kazmierczak, J. J. (1997). Spatial analysis of the distribution of Lyme disease in Wisconsin. *Amercian Journal of Epidemiology,* **145,** 558-665.

Klayman, D. L. (1985). Qinghaosu (artemisinin): an antimalarial drug from China. *Science,* 228,1049-1055.

Kobayashi, J., Vannachone, B., Sato, Y., Sinjo, M., Nambanya, S., Manivang, K. & Inthakone, S. (1998). Current status of malaria infection in a southeastern province of Lao PDR. *Southeast Asian Journal of Tropical Medicine and Public Health*, **29,** 236-241

Kondrashin, A. V., Jun, R. K. & Akiyama. J. (1991) Ecological aspect of forest malaria in southeast Asia. In *Forest malaria in Southesat Asia*, Sharma, V. P. & Kondrashin, A. V (editor). New Delhi.

Kondrashin, A. V. & Rooney. W. (1991). Epidemiology and its control in countries of the WHO Southeast Asia region. *Southeast Asian Journal of Tropical Medicine and Public Health,* **23,** Suppl. 13 – 22.

Kreft, I. & Leeuw, J. D. (1998). *Introducing multilevel modelling.* London: GAGE Publication.

Lacey, L. A. & Lacey C. M. (1990). The medical importance of rice land mosquitoes and their control using alternatives to chemical insecticides. *Journal of American Mosquito Control Association*. **6**, *suppl*. 2-92.

le Sueur, D. L., Ngxongo, S., Sharp, B., Martin, C., Fraser, C., Teuschner, M., Tollman, S., Green, C., Tsoka, J., Solarsh, G. & Mmzaza, A. (1997). *Towards a spatial rural information system*. South Africa: The Health systems trust and medical research council of South African.

le Sueur, D. L., Binka, F., Lengeler, C., de Savigny, D., Snow, R. W., Teuscher, T. & Touré, Y.T. (1997). An atlas of malaria in Africa. *Africa Health,* **19,** 23-24.

Lengeler, C. (1998). Insecticide treated bednets and curtains for malaria control. *The Cochrane Library,* **4,** 1-43

Letham, L. (1996). *GPS made easy, using global positioning systems in the outdoors.* Calgary: Rocky Mountain Book.

Li, X. L., Yu,W. X., Wei, D. R., Li, W., Xu, F., Yang, J. & Zhang, C. Y. (1987). Malaria survey in moving population in Da Luo. *Bulletin of the Yunnan Institute of Malaria Prevention and Treatment.* Vol. 1987 - 1988, 40-42.

Lian. (1984). Epidemiological survey of *P. vivax* malaria in Weisi County, Yunnan. *Chinese Journal of Epidemiology.* 5, 13-19.

Liang, K. Y. & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika,* **73,** 13-22

Lillesand, T. M. & Kiefer, R. W. (1994). *Remote sensing and image interpretation.* New York: John Wiley & Sons, Inc.

Ling, L. C., Liu, K. B. & Yao, Y. T. (1936). Studies on the so-called *changch'i* part II. *Changch'i* in Yunnan. *Chinese Medical Journal,* **50,** 1815-1828.

Lindsay, S. W. & Gibson, M. E. (1988). Bednets revisted-old idea new angle. *Parasitology Today,* 4, 270-272.

Lindsay, S. W., Shenton, F. C., Snow, R. W. & Greenwood, B. M. (1989). Responses of *Anopheles gambiae* complex mosquitoes to the use of untreated bednets in The Gambia. *Medical and Veterinary Entomology,* **3,** 253-262.

Lindsay, S. W., Schellenberg, J. A. R., Zeiler, H. A., Daly, R. J., Salum, F. M. & Wilkins, H. A. (1995). Exposure of Gambian children to *Anopheles gambiae* malaria vectors in an irrigated rice production area. *Medical and Veterinary Entomology,* **9,** 50-58.

Lindsay, S. W., Parson, L. & Thomas, C. J. (1998a). Mapping the ranges and relative abundance of the two principal African malaria vectors, *Anopheles gambiae sensu stricto* and *An. arabiensis*, using climate data. *Proceeding of the Royal Society London, B Biololgy Science.* **65,** 847-854.

Lindsay, S. W. & Martens, W. J. (1998b). Malaria in the African highlands: past, present and future. *Bulletin of the World Health Organization.* **76,** 33-45.

Lindsay, S. W. & Birley, M. H. (1996). Climate change and malaria transmission. *Annals of Tropical Mededicine & Parasitology.* **90,** 573-88.

Lindsay, S. W. & Thomas, C. J. (2000). Mapping and estimating the population at risk from lymphatic filariasis in Africa. *Transaction of the Royal Society of Tropical Medicine Hygiene.* **94,** 37-45.

Lindsay, S. W., Bodker, R., Malima, R., Msangeni, H. A. & Kisinza, W. (2000). Effect of 1997-98 El Nino on highland malaria in Tanzania. *Lancet,* **355,** 989-990.

Lines, J. D., Myamba, J. & Curtis, C. F. (1987) Experimental hut trials of permethrin-impregnated mosquito nets and eave curtains against malaria vectors in Tanzania. *Medical and Veterinary Entomology,* **1,** 37-51.

Ling, L. C., Liu, K. B. & Yao, Y. T. (1936). Studies on the so-called *changch'i* part II. *changch'i* in Yunan. *Chinese Medical Journal,* **50,** 1815-1828.

Linthicum, K. J., Baily, C. L., Davies, F. G. & Tucker, C. J. (1987). Detection of Rift valley fever viral activity in Kenya by satellite remote sensing imagery. *Science,* **235,** 1656-1659.

Linthicum, K. J., Baily, C. L., Tucker, C. J., Mitchell, T. M., Logan, T. M., Davis, F. G., Kamau, C. W., Thande, P. C. & Wasatch, J. N. (1990). Application of polar-orbiting, meteorological satellite data to detect flooding in Rife Valley Fever virus vector mosquito habitats in Kenya. *Medical and Veterinary Entomology,* **4,** 433-438.

Linthicum, K. J., Baily, C. L., Tucker, C. J., Angleberger, D. R., Cannon, T., Logan, T. M., Gibbs, P. H. & Nickeson. (1991). Towards real-time prediction of Rife Valley fever epidemics in Africa. *Preventive Veterinary Medicine,* **11,** 325-334.

Liu, C. F., Qian, H. L., Tang, L. H., Zheng, X., Gu, Z. C. & Zhu, W. D. (1995). Current malaria stratification in China. *Chinese Journal of Parasitology &Parasitic Diseases,* **13,** 8-12

Lobitz, B., Beck, L., Huq, A., Wood, B., Fuchs, G., Faruque, A. S. & Colwell, R. (2000). Climate and infectious disease: use of remote sensing for detection of Vibrio cholerae by indirect measurement. *Proceeding of the National Academy of Science of the U S A.* **97,**1438-1443

Longford, N. F. (1993). *Random Coefficient Models.* Oxford: Clarendon Press.

Lothrop, H. D. & Reisen, W. K. (1999). A geographical information system to manage mosquito and arbovirus surveillance and control data in the Coachella Valley of California. *Journal of American Mosquito Control Association,* **15,** 299-307.

Luo, D. P., LU, D., Yao, R., Li, P., Huo, X., Li, A. M., Wen, L., Ge, C., Zhang, S. W., Huo, H. R., & Shang, L. Y. (1994a). Alphamethrin-impregnated bed nets for malaria and mosquito control in China. *Transaction of the Royal Society of Tropical Medicine & Hygiene,* **88,** 625-628.

Luo, D. P., Zhang, K. H., Song, J. D., Yao, R. G., Hou, H. R., Liu, B. X., Li, Y., & Wang, Z. (1994b). The effect of DDT spraying and bednets impregnated with pyrethriod insecticide on the incidence of Japanese encephalitis virus infection. *Transaction of the Royal Society of Tropical Medicine & Hygiene,* **88,** 629-631.

Luo, D. P., Shang, L. Y., Liu, X. L. &You, X. C. (1996). A successful control programme for falciparum malaria in Xinyang, China. *Transaction of the Royal Society of Tropical Medicine & Hygiene,* **90,** 100-102.

Ma, L. S. (1942), Malaria survey of Wangting, Yunnan. *National Medical Journal of China*, **27,** 587-591.

MacDonald, G. (1957). *The epidemiology and control of malaria,* London: Oxford University Press.

Malone, J. B., Huh, O. K., Fehler, D. P., Wilson, P. A., Wilemsky, D. E., Holmes, R. A. & Elmagdoub, A. L. (1994). Temperature data from satellite imagery and the distribution of *Schistosoma mansoni* in Egypt. *Amercian Journal of Tropical Medicine and Hygiene,* **50,** 714-722.

Malone, J. B., Abdel-Rahman, M. S. E. L., Bahy, M. M., Huh, O. K., Shafik, M. & Bavia, M. (1997). Geographical information systems and the distribution of *Schistosoma mansoni* in the Nile Delta. *Parasitology Today*, **13,**112-117.

Mather, P. M. (1999). *Computer Processing of Remote-Sening Image, An introduction*, 2nd, West Sussex. John Wiley & Son Ltd

McMichael, A. J., Haines, A., Slooff, R. & Kovats, S. (1996). *Climate change and human health: an assessment prepared by a Task Group on behalf of World Health Organization, the World Meteorological Organization and the United Nations Environment Programme.* World Health Organization, Geneva.

Michael, E. & Bundy, D. A. P. (1997). Global mapping of Lymphatic filariasis. *Parasitology Today*, **13,** 472-476.

Molineaux, L. (1988). The epidemiology of human malaria as an explanation of its distribution, including some implications for its control: In: *Malaria: Principles and Practice,* Wernsdorfer WH & McGregor I (editor), London., pp 913-998. Churchill Livingstone.

Mott, K. E., Nuttall, I., Desjeux, P. & Cattand. P. (1995). New geographical approach to control of some parasitic zoonoses. *Bulletin of the World Health Organization,* **73,** 247-257.

Molineaux, D. H. (1997). Patterns of change in vector-borne diseases. *Annals of Tropical Medicine & Parasitology*, **91,** 827-839.

Moncayo, A. C., Edman, J. D. & Finn, J. T. (2000). Application of geographic information technology in determining risk of eastern equine encephalomyelitis virus transmission. *Journal of American Mosquito Control Association,* **16,** 28-35

Narasimham, M. V. V. L. (1991). Perspectives of forest malaria in India. In: *Forest malaria in Southesat Asia,* Sharma, V. P. & Kondrashin, A. V. (editor). New Delhi.

Nicholson, M. C. & Mather, T. N. (1996). Methods for evaluating lyme disease risk using geographical information systems and geographical analysis. *Journal of Medical Entomology,* **33,** 711-720.

Nobre, F. F., Braga, A. L., Pinheriro, R. S. & Lopes, J. A. D. S. (1997). GISEpi: a simple geographical information system to support public health surveillance and epidemiological investigations. *Computer Methods and Programs in Biomedicine,* **53,** 33-45.

Omumbo J. A., Ouma, J., Rapuoda, B., Craig, M. H., le Sueur, D. & Snow. R. W. (1997). Mapping malaria transmission intensity using geographical information systems (GIS); an example from Kenya. *Annals of Tropical Medicine and Parasitology,* **92,** 7-21.

Onori, E. & Grab, B. (1980). Indicators for the forecasting of malaria epidemics. *Bulletin of the World Health Organisation,* **58,** 91-98.

Openshaw, S. (1996). Geographical information systems and tropical disease. *Transaction of the Royal Society of Tropical Medicine and Hygiene,* **90,** 337-339.

Parkinson, C. L. (1997). *Earth from above,* Sausalito, Califonia. University Science Book.

Patz, J. A., Epstein, P. R., Burke, T. A. & Balbus, J. M. (1996). Global climate change and emerging infectious diseases. *JAMA,* **275,** 217-223

Patz, J. A., Martens, W. J. M., Focks, D. A., Jetten, T. H. (1998). Dengue fever epidemic potential as projected by general circulation models of global climate change. *Environmental Health Perspective,* **106,**147-153.

Pavlovsky, E. N. (1945). The ecological Parasitology. *Journal of General Biology,* **6,** 65-92.

Pavlovsky, E. N. (1966). *The natural nidality of transmittable diseases, with special reference to the landscape epidemiology of zooanthroponoses,* Levine, N. D. (editor), Urbana: University of Illionis Press.

Pendergast, J. F., Gange, S. J., Newtow, M. A., Lindstrom, M. J., Palta, M. & Fisher, M. R. (1996). A survey of method for analysis clustered binary response data. *International Statistical Review,* **64,** 89-118.

Pikhart, H., Prikazsky, M., Bobak, M., Kriz, B., Celko, M., Danova, M., Pryl, K. & Pretel, J. (1997). Association between ambient air concentrations of nitrogen dixoxine and respiratory symptoms in children in Frague, Czech republic, preliminary results from the Czech part of the saviah study. *Central European Journal of Public Health,* **5,** 82-85.

P.L.A. Kuanmin medical research institute, malaria research group. (1978). Clinical observation of the sensitivity of *Plasmodium falciparum* to chloroquine in Mendin, Yunnan province. *Chinese Internal Medical Journal,* **18,** 247-249.

Poolsuwan, S. (1995). Malaria in prehistoric Southeast Asia. *Southeast Asian Journal of Tropical Medicine and Public Health,* **26**, 3-22.

Pope, K. O., Sheffner, E. J., Kasischke, E. S., Birney, K., Njogu, A. R. & Roberts, C. R. (1992). Identification of central Kenyan Rift valley fever virus vector habitats with Landsat TM and evaluation of their flooding status with airborne imaging radar. *Remote Sensing and Environment,* **40,** 185-196.

Randolph, S. E. (1993). Climate, satellite imagery and the seasonal abundance of the tick *Rhipicephalus appendiculatus* in southern Africa: a new perspective. *Medical and Veterinary Entomology,* **7,** 243-258.

Rejmankova, E., Robert, D. R., Pawley, A., Manguin, S. & Polanco, J. (1995). Prediction of adult *Anopheles albimanus* density in villages based on distance to remotely sensed larval habitats. *Amercian Journal of Tropical Medicine and Hygiene,* **51,** 482 - 488.

Rejmankova, E., Pope, K. O., Roberts, D. R., Lege, M. G., Andre, R., Greico, J. & Alonzo, Y. (1998).Characterization and detection of *Anopheles vestitipennis* and *Anopheles punctimacula (Diptera: Culicidae)* larval habitats in Belize with field survey and SPOT satellite imagery. *Journal Vector Ecology,* **23,** 74-88.

Ribeiro, J. M. C., Seulu, F., Abose, T., Kidane, G. & Teklehaimanot. (1996). Temporal and spatial distribution of anopheline mosquitoes in an Ethiopian village: implication for malaria control strategies. *Bulletin of the World Health Organization,* **74,** 299-305.

Richard, F. O. (1993). Use of geographic information systems in control programs for onchocerciasis in Guatemala. *Bulletin of Pan American Health Organization,* **27,** 52-55.

Ritchie. S. A. (1993). Application of radar rainfall estimates for surveillance of *Aedes Taeniorhynchus* larvae. *Journal of the American Mosquito Control Association,* **9,** 228-231.

Robertson, R. C. & Chang, T. L. (1940). Malaria Survey in Western Yunnan, Lungling area and Lushih County. *Chinese Medical Journal,* **58,** 446-455.

Robertson, R. C. (1940). Malaria in Western Yunnan with reference to China-Myanmar highway. *Chinese Medical Journal,* **57,** 57-73.

Roberts, D. R. & Rodriguez, M. H. (1994). The environment, remote sensing, and malaria control. *Annals of the New York Academy of Sciences,* 397-402.

Roberts, D. R., Paris, J. F., Manguin, S., Harbach, R. E., Woodruff, R., Rejmankova, E., Polanco, J., Wullschleger, B. & Legters, I. J. (1996). Predictions of malaria vector distribution in Belize based on multispectral satellite data. *Amercian Journal of Tropical Medicine and Hygiene,* **54,** 304-308.

Robert, V., Gazin, P., Boudin, C., Molez, J. F., Ouedraogo, V. & Carnevale, P. (1985). La transmission du zone de savane arboree et en zone rizicole des environs de Bobo-Dioulasso (Burkina Faso). *Annales de la Societe Belge de Medecine Tropicale,* **65** suppl., 201-214

Robert, V., Carnevale, P ., Ouedraogo, V ., Petrarca V & Coluzzi.M. (1988). La transmission du paludisme humain dans un village de savane du sud ouest du Burkina Faso. *Annales de la Societe Belge de Medecine Tropicale,* **68,** 107-121.

Robert, V. (1989). La transmission du paludisme humain: la zone des savanes d'Afrique de l'ouest. These de science. University of Paris VI. France.

Roberts, D. R., Paris, J. F., Manguin, S., Harbach, R. E., Woodruff, R., Rejmankova, E., Polanco, J., Wullschleger, B. & Legters, I. J. (1996). Predictions of malaria vector distribution in Belize based on multispectral satellite data. *Amercian Journal of Tropical Medicine and Hygiene,* **54,** 304-308.

Robinson, T., Roger, D. & Williams, B. (1997a). Univariate analysis of tsetse habitat in the common fly belt of Southern Africa using climate and remotely sensed vegetation data. *Medical and Veterinary Entomology,* **11,** 223-234.

Robinson, T., Roger, D. & Williams, B. (1997b). Mapping tsetse habitat suitability in the common fly belt of southern Africa using multivariate analysis of climate and remotely sensed vegetation data. *Medical and Veterinary Entomology,* **11,** 235-45.

Rodriguez, A. D., Rodriguez, M. H., Hernandez, J. E., Dister, S. W., Beck, L. R., Rejmankova, E. & Roberts, D. R. (1996). Landscape surrounding human settlements and *Anopleles albimanus (Diptera: Culicidae)* abundance in Southern Chiapas, Mexico. *Journal of Medical Entomology,* **33,** 39-48.

Rogers, D. J. & Randolph, S. E. (1991). Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature,* **351,** 739-741.

Rogers, D. J. & Randolph, S. E. (1993). Distribution of Tsetse and Ticks in Africa: past, present and future. *Parasitiology Today,* **9,** 266-271

Ross, R. (1911). Malaria in the community. *The prevention of malaria,* In: R Ross (ed). London; 151-253

Russell, P. F., West, L. S., Manwell, R. D. & MacDonald, G. (1963). *Practical malariaology.* London: Oxford University Press.

Samarsinghe, M. U. L. P. (1986). The present malaria situations in Sri Lanka with particular reference to areas where irrigation has recently been introduced. Workshop on *irrigation and vector borne disease transmission, Digama, Sri Lanka.* Published by the International Irrigation Management Institute.

Schneider, K. & Robbins, P. (1998). *GIS and Mountain Environment, Exploration in Geographic Information Systems Technology* Vol.5 United Nation Institute for Training & Research

Service, M. W. (1991). Agriculture development and anthropod-borne diseases: a review *Rev. Saude Publ. S. Paulo,* **25,** 165-178.

Service, M. W. (1989). Rice, a challenge to health. *Parasitology Today.* **5,** 162-165.

Shang, L. Y. & Hou, G. P. (1992). The prevalent features and control of malaria in Henan. *Chinese Journal of Parasitic Disease Control*, **5**, 1-3

Shaw, G. & Wheeler, D. (1994). *Statistical techniques in geographical analysis*. London: David Fulton Publishers.

Sharma, V. P., Dhiman, R. C., Ansari, M. A., Nagpal, B. N., Srivastava, A., Kmanavalan, P., Adiga, S., Radhakrishinan, K. & Chandrasekhar, M. G. (1996). Study on the feasibility of delineating mosquito genic conditions in and around Delhi using India Remote Sensing Satellite data. *Indian Journal of Malariology*, **33**, 107-125.

Sharma, V. P., Nagpal, B. N., Srivastava, A., Adiga, S. & Manavalan, P. (1996). Estimation of larval production in Sanjay Lake and its surrounding ponds in Delhi, India using remote sensing technology. *Southeast Asian Journal of Tropical Medicine and Public Health*, **27**, 834-840.

Sharma, V. P. & Srivastava, A. (1997). Role of geographic information system in malaria control. *Indian Journal of Medical Research*, **106**, 198-204.

Sharma, V. P. & Kondrashin, A. V. (1991). *Forest Malaria in Southeast Asia*. WHO/MRC, Delhi

Singh, G. P., Chitkara, S., Kalra, N. L., Makepur, K. B. & Narasimham, M. V. (1990). Development of a methodology for malariogenic stratification as a tool for malaria control. *Journal Communicable Diseases*, **22**, 1-11

Skole, D. & Tucker, C. (1993). Tropical Deforestation and Habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science*, **260**, 1905-1910

Snow, J. (1849). *On the mode of communication of cholera*. 2nd, London: J. Churchill,.

Snow, R. W., Rowan, K. M., Lindsay, S. W. & Greenwood, B. M. (1988). A trial of bed nets (mosquito nets) as a malaria control strategy in a rural area of The Gambia, West Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **82**, 212-215.

Snow, R.W., Menon, A. & Greenwood, B. M. (1989). Measuring morbidity from malaria, *Annals of Tropical Medicine and Parasitology*, **83**, 321-323

Snow R. W., Marsh, K. & le Sueur. D. (1996). The need for maps of transmission intensity to guide malaria control in Africa. *Parasitology Today*, **12**, 455-457.

Snow R. W., Gouws, E., Omumbo, J. A., Rapuoda, B., Craig, M. H., Tanser, F. C., le Sueur, D. & Ouma. J. (1998). Models to predict the intensity of Plasmodium falciparum transmission: applications to the burden of disease in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **92**, 601-606.

Snow R. W., Craig, M. H., Deichmann, U. & le Sueur, D. (1999a). A continental risk map for malaria mortality among African children. *Parasitology Today*, **15**, 99-104.

Snow R. W., Craig, M. H., Deichmann, U., Marsh, K. & Lopez, A. (1999b). Estimating mortality, morbidity and disability due tomalaria among Africa's non-pregnant population. *Bulletin of the World Health Organization,* **77,** 624-640.

Swaroop, S. (1946). Forecasting of epidemic malaria in Punjab, India. *American Journal of Tropical Medicine,* **29,** 1-17.

Tang, L. H., Qian, H. L. & Xu, S. H. (1991). Malaria and its control in the People's Republic of China. *Southeast Asian Journal of Tropical Medicine and Public Health,* **22,** 467-476.

Thomas, C. J. & Lindsay, S. W. (2000). Local scale variation in malaria infection among rural Gambian children estimated by satellite remote sensing. *Transaction of the Royal Society of Tropical Medicine and Hygiene,* **94,** 159-163.

Thomson, M. C., D'Alessandro, U., Bennett, S., Connor, S. J., Langerock, P., Jawara, M., Todd, J. & Greenwood, B. M. (1994). Malaria prevalence is inversely related to vector density in The Gambia, West Africa. *Transaction of the Royal Society of Tropical Medicine and Hygiene,* **88,** 638-643.

Thomson, M. C., Connor, S. J., Milligan, P. J. & Flasse, S. P. (1996). The ecology of malaria--as seen from Earth-observation satellites. *Annals of Tropical Medicine and Parasitology,* **90,** 243-264.

Thomson, M. C., Connor, S. J., Milligan, P. J. & Flasse, S. P. (1997). Mapping malaria risk in Africa: what can satellite data contribute? *Parasitology Today,* **13,** 313-318.

Thomson, M. C., Elnaiem, D. A., Ashford, R. W. & Connor, S. J. (1999). Towards a kala azar risk map for Sudan: mapping the potential distribution of *Phlebotomus orientalis* using digital data of environmental variables. *Tropical Medicine and International Health,* **4,**105-113.

Thomson, M. C., Connor, S. J., D'Alessandro, U., Rowlingson, B., Diggle, P., Cresswell, M. & Greenwood, B. (1999). Predicting malaria infection in Gambian children from satellite data and bed net use surveys: the importance of spatial correlation in the interpretation of results. *Amercian Journal of Tropical Medicine and Hygiene,* **61,** 2-8.

Thompson, R., Begtrup, K., Cuamba, N., Dgedge, M., Mendis, C., Gamage-Mendis, A., Enosse, S. M., Barreto, J., Sinden, R. E. & Hogh, B. (1997). The matola malaria project: a temporal and spatial study of malaria transmission and disease in a suburban area of Maputo, Mozambique. *Amercian Journal of Tropical Medicine and Hygiene,* **57,** 550-559.

Tin, F. & Tun, T. (1991). Forest related malaria in Myanmar. In: *Forest malaria in Southesat Asia.* Sharma, V. P. & Kondrashin, A. V(editor). New Delhi.

Wagner, V. E., Rowley, R. H., Narlock, S. A. & Newson. H. D. (1979). Remote sensing: a rapid and accurate method of data acquisition for a newly formed mosquito control district. *Mosquito News,* **39,** 283-287.

Walsh, J. F., Molyneux, D. H. & Birley, M. H. (1993). Deforestation: effects on vector-borne disease. *Parasitology,* **106,** S55-S75.

Walton, G. A. (1947). On the control of malaria in Freetown, Sierra Leone: 1. *Plasmodium falciparum* and *Anopheles gambiae* in relation to malaria occuring in infants. *Annals of Tropical Medicine and Parasitology*, **41**, 380-470

Wang, S. W., He, B., Hou, X. W. & Cen, R. P. (1957). *Summary of Malaria Control and Investigation In Yunnan Military Forces*, Beijing: People's Military Press.

Wang, Y. (1990). *Yunnan Agricultural Meteorological Resources and its Stratification*, Beijing: Meteorological Press.

Welch, J. B., Olson, J. K., Hart, W. G., Ingle, S. G. & Davis, M. R. (1989). Use of aerial color IR as survey technique for *Psorphora columbiae* oviposition habitats in Texas (USA) ricelands. *Journal of the American Mosquito Control Association*, **5**, 147-160.

Ward, M. H., Nuckols, J. R., Weigel, S. J., Maxwell, S. K., Cantor, K. P. & Miller, R. S. (2000). Identifying populations potentially exposed to agricultural pesticides using remote sensing and a Geographic Information System. *Environmental Health Perspective*, **108**, 5-12.

Washino, R. K. & Wood, B. L. (1994). Application of remote sensing to arthropod vector surveillance and control. *Amercian Journal of Tropical Medicine and Hygiene*, **50** (suppl), 134-144.

Welch, J. B., Olson, J. K., Yates, M. M., Benton. A. R. & Barker, R. D. (1989). Conceptual model for the use of aerial color infrared photography by districts as a survey technique for *Psorphora columbiae* oviposition habitats in Texas ricelands. *Journal of the American Mosquito Control Association*, **5**, 369-373.

Wernsdorfer, W. H. (1991). The development and spread of drug-resistant malaria. *Parasitology Today*, **7**, 295-303.

Wernsdorfer, W. H., Chongsuphajiasiddhi, T. & Salazar, N. P. (1994). A symposium on containment of mefloquine-resistant falciparum malaria in Southeast Asia with special reference to border malaria. *Southeast Asian Journal of Tropical Medicine and Public Health*, **25**, 11-18.

WHO. (1957). *Expert Committee on malaria: six report*. Techical Reported series, **123**, 59-62.

WHO. (1963). *Terminology of Malaria and Malaria Eradication*, Geneva

WHO (1986). *WHO Expert Committee on Malaria: eighteenth report*. Technical report series, 735.

WHO. (1988). Malaria diagnosis: memorandum from a WHO meeting. *Bulletin of the World Health Organization*, **65**, 575-594.

WHO. (1993). Global malaria control. *Bulletin of the World health Organization*, **71**, 281-286.

WHO. (1996). World malaria situation in 1993. *Weekly Epidemiological Record*, **71**, 41-48.

Williams, J. I. (1941). Malaria on the China Myanmar highway. *American Journal of Tropical Medicine and Hygiene*, **21**, 1-11.

Wood, B. L., Beck, L. R., Washino, R. K., Palchick, S. M. & Sebesta, P. D. (1991). Spectral and spatial characterization of rice field mosquito habitat. *International Journal of Remote Sensing*, **12**, 621-626.

Wood, B. L., Beck, L. R., Washino, R. K., Hibbard, K. A. & Salute, J. S. (1992). Estimating high mosquito-producing rice fields using spectral and spatial data. *International Journal of Remote Sensing*, **13**, 2813 – 2826.

Xu, J. W. & Liu, H. (1997). Border malaria in Yunnan, China. *Southeast Asian Journal of Tropical Medicine and Public Health,* **28**, 456-459.

Yang, C. (1991). Malaria prevalence in Yunnan. *Yunnan Di Fang Bing Zi Liao Hui Bian*, **4**, 18-22.

Yang, C., Yang, P. F., Gao, H. P., Sheng, Y. M. & Li, A. F. (1992). A survey of *P. falciparum* malaria outbreak in Nujiezai Township, Ailaoshan, Yunnan. *Bulletin of the Yunnan Institute of Malaria Prevention and Treatment.* Vol. 1991 - 1992, 31-34.

Yang, H. L. (1989). A field study on the microtechnique method for susceptibility of *Plasmodium falciparum* to pyronaridine and artesunate. *Chinese Journal of Parasitic Diseases*, **2**, 169-171.

Yang, H., Xu, J. W., Wang, W. Z., Ma, J., Yang, G. C., Guo, Y. A., Wang, L. B., Liu, X. Z., Yang, Z, C., Ma, X. W. Yang, X. W. & Yang, J. C. (1999). Current malaria control strategy and their evaluation in Red River basin, Yunnan. *Endemic Disease Bulletin*, **14**, 56-58.

Yang, H. L., Yang, P. F. & Yang, Y. M. (1994). *In vitro* sensitivity of *Plasmodium falciparum* to mefoquine, quinine, amodiaquine, chloroquine, pyronaridine and sulfadoxine/pyrimethamine in south Yunnan. *Chinese Journal of Parasitology & Parasitic Diseases,* **12**,140-142.

Yang, Y. M. & Li, C. F. (1992). A survey of malaria outbreaks in a new rubber plantation, Yunnan. *Bulletin of the Yunnan Institute of Malaria Prevention and Treatment.* Vol. 1991 - 1992, 111 –112.

Yao, Y. T. & Pei, Y. S. (1943). Some epidemiological factors of malaria in Mangshih, Yunan, with remarks on occurrence of blackwater fever. *Chinese Medical Journal*, **61**, 197-211.

Zhang, X. K. (1990). Analysis of geographical distribution of malaria in Yunnan. *Yunnan Medicine*, **10**, 291-295.

Zhang, Z. X. (1992). Investigation on malaria situation and control in the temporary resident area of people going abroad for farming and returning to Tenchong, Yunnan. *Bulletin of the Yunnan Institute of Malaria Prevention and Treatment* Vol. 1991 - 1992, 105 –107.

Zhang, Z. X. & Yang, C. J. (1996). Application of deltamethrin-impregnated bednets for mosquito and malaria control in Yunnan, China. *Southeast Asian Journal of Tropical Medicine and Public Health,* **27**. 367-371.

Zhao, S. L. (1993). Malaria vertical stratification in Yunnan. *Yunnan Medicine*, **14**, 296 - 300.

Zhou, Z. J. (1991). *Malaria Control and Research in China*, (editor), Beijing, pp. 166-168. People's Health Press.

Zhou, Z. J. (1981). The malaria situation in the People's Republic of China, *Bulletin of the World Health Organization*, **59**, 931 – 936.

Zhu, D. F. (1994). Prevalent outline of falciparum malaria and the resistant of *P. falciparum* to chloroquine in Yunnan Province. *Chinese Journal of Parasitic Disease Control*, **7**, 84-87.

Zhu, D. F., Che, L. G. & Su, F. C. (1994). The malaria situation on the frontiers of Yunnan Province, China. *Southeast Asian Journal of Tropical Medicine and Public Health*, **25**, 19-24.

Appendix I. Questionnaires

Form 1. Census form

Name/ID

Name of administrative village/Code No.

Name of village/Village code

Name of household/Household code

Sex (male=0 female=1)

Date of birth

Ethnic group

Education
1 Illiterate
2 Elementary school
3 Junior high school
4 Senior high school

Do you use mosquito nets (yes=1, no=0)

Form 2. Geographical location recording form

Name of household/Household code

Name of village/Village code

Name of administrative village/Code No.

Geographical location of household by GPS

Nature of house
1 Temporal hut
2 Built with earth brick and straw roof
3 Built with cement bricks or regular board

Form 3. Febrile patient recording form

Name/ID

Name of administrative village/Code No.

Name of village/Village code

Name of household/Household code

Sex (male=0 female=1)

Date of birth

Date of first febrile episode in last 7 days

Did you visit (work) in the low altitude area in the last month (yes=1 no=0)

If yes, when did you start =visiting the lowland

Date of visit

Code No. of blood slides

Method of slide collection (active=1, passive=2)

Form 4. Malaria case recording form

Name/ID

Name of administrative village/Code No.

Name of village/Village code

Name of household/Household code

Sex (male=0 female=1)

Date of birth

Date of slide collection

Date of slide examination

Malaria parasite (Negative=0, *P. vivax*=1, *P. falciparum*=2, mixed=3)

Laboratory Techinician

Feature of infections (indigenous=1, imported = 0)

Appendix II

Location of GCPs on image and in UTM co-ordinates, with RMS error

GCPs	Input x	Input y	Output x	Output y
1	162.188	-114.813	302086.168	2558930.187
2	227.313	-129.188	303345.982	2558436.703
3	314.813	-189.313	304851.896	2556952.332
4	762.375	-203.875	313582.288	2555059.569
5	762.125	-249.125	313425.992	2554166.267
6	834.863	-288.063	314745.500	2553149.964
7	968.625	-273.875	317429.942	2552932.353
8	387.875	-365.875	305696.592	2553184.721
9	596.875	-361.125	309830.554	2552562.440
10	653.156	-386.906	310826.073	2551820.595
11	375.125	-695.125	304317.374	2546750.790
12	98.625	-820.125	298441.986	2545265.520
13	319.938	-903.938	302508.257	2542799.322
14	309.938	-1007.938	301958.309	2540840.519
15	377.438	-1066.938	303059.414	2539419.951
16	355.031	-587.031	304314.338	2548919.880
17	768.656	-500.031	312740.445	2549190.756
18	814.781	-579.781	313404.653	2547422.668

RMS error 0.834