

Malaria Risk Mapping and Prediction in Côte d'Ivoire

Charles M. van Wijk



*Thesis presented in partial fulfilment of the requirements for the degree of
Master of Natural Science at the University of Stellenbosch*

*Study leaders: Prof HL Zietsman & Dr T Teuscher
December 1999*

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

OPSOMMING

Malaria is tans die belangrikste en dodelikste insekkoordraagbare siekte in Afrika, wat 'n al hoe groter wordende gesondheids- en ekonomiese impak laat in die ontwikkelende wêreld. Aangesien die verspreiding en intensiteit van malaria beïnvloed word deur 'n groot aantal omgewingsfaktore, kan mediese geografie – die studie van siektes se ruimtelike mens-omgewing interverwantskappe – 'n beduidende rol speel ter voorspelling van gesondheidsrisiko's in verskillende geografiese areas, soos gebaseer op omgewingskenmerke. Die ondersoek het dus die volgende ingesluit: die ontwikkeling van malaria- risikokaarte vir die noordelike Ivoorkus studiegebied, 'n ontleding van ruimtelike variasies deur eksploratiewe ruimtelike data-analise en Geografiese Inligtingstelsel (GIS) funksionaliteit, en die soektog na betroubare voorspellingsmodelle soos gebaseer op omgewingsfaktore.

Die intensiteit van malaria-oordrag deur die *Anopheles funestus* en *Anopheles gambiae* muskietspesies is bestudeer as faktore wat 'n direkte invloed het op die malaria-oordrag siklus, terwyl indirekte omgewingsfaktore wat bestudeer is ingesluit het: vleiland-rysverbouing, waterliggame en hul vloedingspatrone, temperatuur- en relatiewe humiditeitsgrense, hoeveelheid reënval en sonskyn, afstande tussen geskikte broeigebiede en menslike woongebiede, en plantegroei digtheid. Malaria en omgewingsdata is verkry vanuit voortgesette opnames gedoen deur Institut Pierre Richet/OCCGE en die WARDA Gesondheidsnavorsing Konsortium in 'n steekproef van 24 noord-Ivoorkus woongebiede, distriksvlak meteorologiese observasies, en GIS en satellietbeeld verwerkings. Ingesamelde data is gemanipuleer, gestandaardiseer en verwerk tot drie indikatore van malariarisiko, en vyf onverwante omgewingsfaktore soos bepaal deur hoofkomponent-analise.

Die toepassing van eksploratiewe ruimtelike data-analise, waarin die ruimtelike funksionaliteit van GIS 'n beduidende rol gespeel het, het behels:

- a) Die kartering van die siekte se ruimtelike verspreiding en van hoërisikogebiede deur middel van proporsionele puntsimbool- en waarskynlikheidskaarte.

- b) Die ondersoek en modellering van eerste- en tweede-orde ruimtelike siektevariasies deur getrianguleerde onreëlmatige netwerk model (TIN) interpolasie, semi-variogramme en oppervlakspatroom-analise.
- c) Die vasstelling van sigbare en statistiese verhoudings tussen malariavoorkoms en omgewingsfaktore, deur proporsionele puntsimboolkaarte en meervoudige regressie-analise, in die soeke na betroubare voorspellingsmodelle. Residuele van finale modelle is gekarteer om moontlike patrone in oor- en onderskatting te ondersoek.

'n Algemene toename in siektevoorkoms en -oordrag is gedurende die navorsingsperiode geïdentifiseer. Areas met 'n hoë malariarisiko is bepaal, en die ondersoek tot eerste-orde ruimtelike variasies het hierdie hoërisikogebiede beklemtoon. Semi-variogramme het 'n hoë mate van ruimtelike variasie getoon, met min, indien enige, ruimtelike verband tussen observasies. Geen duidelike tweede-orde variasie kon dus bepaal word nie, asook geen aanvaarbare eerste- of tweede-orde ruimtelike variasiemodel afgelei word nie. Relatiewe humiditeit- en temperatuurmaatstawwe is visueel geïdentifiseer as moontlike omgewingsfaktore wat die voorkoms van malaria kan beïnvloed, en is ook statisties deur regressiemodellering bevestig om 'n relatief groot persentasie van siektewisseling te verklaar, tesame met metings van afstande en plantegroeidigheid. Die lae vlak van variansie verklaar deur sommige modelle, tesame met oneweredige ruimtelike verspreiding en hoë waardes van residuele, het betroubare voorspellings egter onmoontlik gemaak. Dit is derhalwe duidelik dat die verhoudings tussen malariavoorkoms of -oordrag en omgewingsfaktore kompleks, en dikwels tot spesifieke kombinasies in 'n area beperk is. Dit benodig waarskynlik hoër-orde polinomiale funksies en 'n wyer spektrum omgewingsdeterminante, wat nog geïdentifiseer en in opvolgstudies gebruik kan word.

SUMMARY

Malaria, being the most important vector-borne disease in Africa, is still leaving an ever-increasing trail of health and economic impediments within the developing world. Since malaria's spatial distribution and intensity are influenced by numerous environmental variables, medical geography, the study of the spatial human-environmental interrelationship of disease, can make a significant contribution to modelling disease risk for different geographic locations, as based on environmental determinants. The investigation consequently included the development of malaria risk maps within northern Côte d'Ivoire, an analysis of its spatial variation through exploratory spatial data analysis and Geographical Information System (GIS) functionality, and the search for predictive models using environmental variables.

Malaria transmission intensities of the *Anopheles funestus* and *Anopheles gambiae* mosquito species have been studied as factors having a direct influence on the malaria transmission cycle, whilst indirect environmental influencing factors which were studied included: wetland rice production, water bodies and their flooding pattern, temperature and relative humidity ranges, amount of rainfall and insolation, distances between aquatic breeding grounds and villages, and vegetation cover. Data have been obtained from ongoing surveys carried out by Institut Pierre Richet/OCCGE and the WARDA Health Research Consortium within a random selection of 24 villages in northern Côte d'Ivoire, district level meteorological observations, and GIS and satellite imagery computations. The data sets have been manipulated, standardised and reduced to three malaria risk indicators, and five uncorrelated environmental measures through principal component analysis.

The application of exploratory spatial data analysis, in which the spatial capabilities of GIS played a significant role, involved:

- a) Mapping the disease's spatial distribution and extreme risk areas by means of proportional circle and probability maps.

- b) Exploring and modelling first and second order spatial disease variation through Triangulated Irregular Network (TIN) interpolation, semi-variograms, and trend surface analysis.
- c) Visual and statistical detection of associations between malaria and environmental determinants, by means of proportional circle maps and multiple regression analysis, in the search for predictive models. Residuals of final models have been mapped to investigate possible patterns in over- and underestimation.

A general increase in disease incidence and transmission levels during the study period, as well as areas of high risk have been identified. Investigation of first order variation highlighted these high risk areas. Semi-variograms revealed a high degree of spatial variability, with little or no spatial dependence between observations. No clear second order variation could therefore be identified, nor any reasonable first or second order spatial variation model be deduced. Relative humidity and temperature measures appear visually to be associated with malaria occurrence, and were statistically confirmed to account for a substantial percentage of disease variance through regression modeling, together with distance and vegetation density measures. However, the small proportion of variance explained by some models, together with the irregular spatial distribution and high levels of residuals, made reliable prediction impossible. It is therefore clear that the relationships between malaria morbidity or transmission and the environmental factors are complex and often highly site specific, possibly requiring higher order polynomial functions and a wide spectrum of determinants, still to be identified and included in subsequent studies.

ACKNOWLEDGEMENTS

The financial assistance of the Centre for Science Development (HSRC, South Africa) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the Centre of Science Development.

This thesis has been supported by the WARDA/WHO-PEEM/IDRC/DANIDA/ Government of Norway Consortium Research Project on the Association between Irrigated Rice Ecosystems and Vector-borne diseases in West Africa.

CONTENTS

DECLARATION	i
OPSOMMING	ii
SUMMARY	iv
ACKNOWLEDGEMENTS	vi
CONTENTS	vii
FIGURES	ix
TABLES	ix
CONCEPTS AND ACRONYMS	x
CHAPTER 1: A GEOGRAPHICAL CONTRIBUTION TO MALARIA RESEARCH	1
1.1 MALARIA – THE RAMPANT KILLER.....	1
1.2 MALARIA’S HEALTH IMPACT.....	1
1.3 MALARIA’S ECONOMIC IMPACT	2
1.4 THE CONTRIBUTION OF GEOGRAPHY AND GIS IN STUDYING MALARIA	2
1.5 AIMS AND OBJECTIVES	4
1.6 AN OVERVIEW OF THE STUDY AREA.....	6
1.7 CONCLUSION	8
CHAPTER 2: MALARIA AT A GLANCE	10
2.1 THE MALARIA CYCLE	10
2.2 THE MALARIA VECTOR.....	11
2.3 POSSIBLE FACTORS INFLUENCING THE MALARIA CYCLE	12
2.3.1 MALARIA TRANSMISSION INTENSITY	13
2.3.2 WETLAND RICE PRODUCTION’S INFLUENCE.....	14
2.3.3 OTHER POSSIBLE INFLUENCING FACTORS	17
2.3.3.1 <i>Water Bodies and Flooding Pattern</i>	17
2.3.3.2 <i>Temperature</i>	17
2.3.3.3 <i>Relative Humidity</i>	17
2.3.3.4 <i>Rainfall</i>	17
2.3.3.5 <i>Insolation</i>	18
2.3.3.6 <i>Distance from Village to Water Body</i>	18
2.3.3.7 <i>Vegetation</i>	19
2.4 CONCLUSION	19

CHAPTER 3: DATA COLLECTION, MANIPULATION AND REDUCTION.....	20
3.1 DATA COLLECTION AND MANIPULATION	20
3.1.1 MALARIA HEALTH RISK MEASURES	20
3.1.2 WATER BODIES AND ITS FLOODING PATTERN	21
3.1.3 METEOROLOGICAL MEASURES.....	22
3.1.4 DISTANCE FROM VILLAGE TO WATER BODY.....	22
3.1.5 VEGETATION BIOMASS MEASURE	22
3.2 DATA REDUCTION.....	24
CHAPTER 4: EXPLORATORY SPATIAL DATA ANALYSIS	26
4.1 MAPPING MALARIA RISK	26
4.1.1 GEOGRAPHICAL DISTRIBUTION OF MALARIA	27
4.1.2 PROBABILITY MAPS OF DISEASE INCIDENCE AND TRANSMISSION LEVELS.....	30
4.2 EXPLORING SPATIAL DISEASE VARIATION	32
4.2.1 EXPLORING FIRST ORDER VARIATION.....	35
4.2.2 EXPLORING SECOND ORDER VARIATION	35
4.2.3 MODELLING FIRST ORDER VARIATION	39
4.3 ASSOCIATIVE ANALYSIS	41
4.3.1 VISUAL IDENTIFICATION OF ASSOCIATIONS	41
4.3.2 STATISTICAL ANALYSIS OF ASSOCIATIONS	46
4.4 CONCLUSIONS	60
CHAPTER 5: CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH	62
<i>LIST OF REFERENCES.....</i>	67
APPENDIX A: FACTOR LOADINGS	76
APPENDIX B: SEMIVARIOGRAMS	77

FIGURES

FIGURE 1: WETLAND RICE CULTIVATION SYSTEM.....	5
FIGURE 2: EXAMPLE OF A TYPICAL VILLAGE INCLUDED IN THE STUDY	7
FIGURE 3: LIFE WITHIN A VILLAGE SELECTED FOR THE STUDY	7
FIGURE 4: THE MOIST SAVANNA OF CENTRAL NORTHERN CÔTE D'IVOIRE.....	8
FIGURE 5: STUDY AREA: NORTHERN CÔTE D'IVOIRE	9
FIGURE 6: LIFE-CYCLE OF THE PLASMODIUM MALARIA PARASITE IN THE HUMAN AND MOSQUITO HOSTS	10
FIGURE 7: RICE PADDIES, SERVING AS EXTENDED VECTOR BREEDING SITES.....	14
FIGURE 8: NDVI FOR SOME STUDY VILLAGES	23
FIGURE 9: FEVER EPISODES: MAR - JUN '97	28
FIGURE 10: MALARIA TRANSMISSION INTENSITY: JAN - JUN '97.....	29
FIGURE 11: FREQUENCY AND NORMAL DISTRIBUTION OF FEVER EPISODES.....	31
FIGURE 12: FREQUENCY AND NORMAL DISTRIBUTION OF <i>ANOPHELES FUNESTUS</i> TRANSMISSION	31
FIGURE 13: FREQUENCY AND NORMAL DISTRIBUTION OF <i>ANOPHELES GAMBIAE</i> TRANSMISSION	32
FIGURE 14: EXTREME FEVER EPISODE VALUES: MAR – JUN '97	33
FIGURE 15: EXTREME MALARIA TRANSMISSION LEVELS: JAN – JUN '97	34
FIGURE 16: FEVER EPISODE INTERPOLATION USING TIN	36
FIGURE 17: <i>ANOPHELES FUNESTUS</i> TRANSMISSION INTERPOLATION USING TIN.....	37
FIGURE 18: <i>ANOPHELES GAMBIAE</i> TRANSMISSION INTERPOLATION USING TIN.....	38
FIGURE 19: FEVER EPISODES AND HUMIDITY MEASURES: MAR - JUN '97.....	42
FIGURE 20: MALARIA TRANSMISSION LEVEL AND HUMIDITY MEASURES: JAN - JUN '97	43
FIGURE 21: FEVER EPISODES AND WATER BODY SIZE MEASURES: MAR - JUN '97	44
FIGURE 22: MALARIA TRANSMISSION LEVEL AND WATER BODY SIZE MEASURES: JAN - JUN '97	45
FIGURE 23: FEVER EPISODES AND TEMPERATURE MEASURES: MAR - JUN '97.....	47
FIGURE 24: MALARIA TRANSMISSION LEVEL AND TEMPERATURE MEASURES: JAN - JUN '97	48
FIGURE 25: FEVER EPISODES AND DISTANCE MEASURES BETWEEN VILLAGE AND WATER BODIES: MAR - JUN '97.....	49
FIGURE 26: MALARIA TRANSMISSION LEVEL AND DISTANCE MEASURES BETWEEN VILLAGE AND WATER BODIES: JAN - JUN '97.....	50
FIGURE 27: FEVER EPISODES AND BARRENNESS MEASURES: MAR - JUN '97	51
FIGURE 28: MALARIA TRANSMISSION LEVEL AND BARRENNESS MEASURES: JAN - JUN '9752	52
FIGURE 29: RESIDUALS OBTAINED FROM THE FEVER EPISODE MODEL.....	57
FIGURE 30: RESIDUALS OBTAINED FROM THE <i>ANOPHELES FUNESTUS</i> MODEL	58
FIGURE 31: RESIDUALS OBTAINED FROM THE <i>ANOPHELES GAMBIAE</i> MODEL	59

TABLES

TABLE 1: METEOROLOGICAL CONDITIONS WITHIN THE STUDY AREA	6
--	---

CONCEPTS AND ACRONYMS

Agro-ecosystem: An area, with homogeneous physical attributes, within which a certain farming system is practiced.

CHAART: Center for Health Applications of Aerospace Related Technologies, NASA Ames Research Center, Moffet Field, California, USA.

Entomological inoculation rate: See “Malaria transmission rate”.

FAO: Food and Agricultural Organization, Rome, Italy.

Geographic Information Systems (GIS) are computer-based systems for automating, manipulating, displaying and analyzing mapped information. It includes spatial data in the form of geographic coverages (map layers) and descriptive information in the form of a relational database associated with the mapped features (Kitron, Pener, Costin, Orshan, Greenberg & Shalom 1994).

IDRC: International Development Research Center, Ottawa, Canada.

Isotropy: Dependence is purely a function of the distance between two observation points, and not of the direction (Bailey & Gatrell 1995:33).

Lowland: Low and relatively level land at a lower elevation than adjoining districts, usually applied to land less than 600 feet above the sea-level (Stamp & Clark 1979).

Malaria transmission intensity: See “Malaria transmission rate”.

Malaria transmission rate: The number of bites inflicted on one individual by mosquitoes whose salivary glands contain sporozoites that are actually infective, during a specified time period.

Multicollinearity: A high correlation exists between the independent variables.

OCCGE: Organisation de Coordination pour la lutte Contre les Grandes Endemies (Organization for the Control and Cooperation in Control of the Major Endemic Diseases, e.g. malaria, schistosomiasis, onchocerciasis, sleeping sickness, meningitis, etc.), Bobo Dioulassou, Burkina Faso.

Parasite load: The number of parasites per microliter blood.

Parasite rate: Presence of any number of parasites in a blood sample (i.e. greater than 0).

Pathogen: Subject which makes you ill (including viruses, bacteria, parasites, but also chemicals, etc.) (Teuscher, 1999).

PEEM: “Joint WHO/FAO/UNEP/UNCHS Panel of Experts on the Environmental Management for Vector Control”, managed by the PEEM Secretariat, WHO, Geneva.

Spatial autocorrelation: The correlation between values of the same variable at different spatial locations (Bailey & Gatrell 1995:269).

Standardization: Each data value is mean centred by subtracting from it the mean of its corresponding variable, and then further scaled by dividing by the standard deviation of the same variable.

Stationarity: The mean and variance are independent of absolute location and constant throughout the study region (Bailey & Gatrell 1995:33).

Swamp: A waterlogged area and the characteristic vegetation which colonizes it (Goudie 1985).

UNCHS: United Nations Center for Human Settlements, Geneva, Switzerland.

UNEP: United Nations Environment Program, Nairobi, Kenya.

Vector-borne disease: Disease transmitted from one organism to the next via an insect (vector), for example certain anopheline mosquito species in the case of malaria.

WARDA: West Africa Rice Development Association, Bouaké, Côte d'Ivoire

Wetlands: Usually flat areas which are covered either periodically, regularly or permanently by shallow water up to a depth of 6m, with a characteristic fauna and flora. They include swamps and marshes, bogs, fens, pans, water sponges, peatlands, seasonally flooded land around lakes and rivers or level, ill-drained plains. The water may be either fresh or brackish (Goudie 1985; Fuggle & Rabie 1992).

WHO: World Health Organization, Geneva

Zoonotic reservoir: An animal which will harbour the pathogen and in which the pathogen can multiply and be picked up again by a new vector (the zoonotic host does not necessarily suffer from the illness) (Teuscher, 1999).

CHAPTER 1: A GEOGRAPHICAL CONTRIBUTION TO MALARIA RESEARCH

1.1 MALARIA – THE RAMPANT KILLER

“Malaria is one of the planet’s deadliest diseases and one of the leading causes of sickness and death in the developing world” (International Development Research Centre 1996). Malaria and AIDS have been declared by the World Health Organization (WHO) to be the only two diseases in the world whose incidences are increasing at a significant rate. Although AIDS is receiving worldwide attention, malaria remains the most important vector-borne disease [i.e., induced by a disease-carrying insect] that affects human morbidity and mortality on the African continent (McMillan & Meltzer 1996:579; Ngxongo 1993). In fact, according to Capraro (1997), the annual number of worldwide malaria related deaths is more than that induced by AIDS during the past 15 years! However, only a fraction of the money annually demarcated for AIDS research is allocated for malaria studies (some R287 million for malaria, in comparison with AIDS’ R4.35 billion).

In view of a malarial research backlog, this study proposes to investigate the nature of the disease’s spatial variation within a malaria stricken area in northern Côte d’Ivoire, by employing an analytical method known as exploratory spatial data analysis, together with the spatial functional capabilities of Geographical Information Systems (GIS). The first chapter consequently aims to sketch a broad background to the seriousness of the problem at hand, especially malaria’s health and economic impediments. It further outlines the role which this study could play within a wider network of malaria research, noting its specific research objectives, and introduces the study area of interest.

1.2 MALARIA’S HEALTH IMPACT

About 40% of the world’s population – approximately two billion people – are at risk of malaria parasite infection in about 90 countries and territories (Stuttaford 1994:28), with an estimated 300-500 million people worldwide who are infected annually (International Development Research Centre 1996). Eighty to ninety percent of infections occur in sub-Saharan Africa, fundamentally contributing to the region’s highest per capita burden of disease

in the world (Medical Research Council 1996; 1997). WHO's annual clinical case estimates for sub-Saharan Africa is placed at 270-480 million, which includes 140-280 million cases in children under five years old. Being most vulnerable to infections and death, it therefore comes as no surprise that of the annual 1.4-2.8 million malarial related deaths in Africa, tragically, one million of these are children under five years of age (McMillan & Meltzer 1996:580; World Health Organization 1992; International Development Research Centre 1996). With more than 2,700 children deaths per day, malaria is therefore the single largest killer of children in Africa (Medical Research Council 1996:1-2; World Bank 1994:18).

1.3 MALARIA'S ECONOMIC IMPACT

In spite of being the first, largest, and still-increasing component of Africa's burden of disease, the malaria situation worldwide is in many places deteriorating, compared with 10 years ago, as malaria parasites become more resistant to chloroquine and other malarial drugs (World Bank 1994:18; Medical Research Council 1997). The effects of poor health, however, go far beyond physical pain and suffering. It contributes to a serious impediment to economic development and productivity, due to pervasive illness of its work force, low life expectancy, compromise of learning, constraint of environments for entrepreneurial and productive activities, and consequently major work-loss (Brêtas 1995: 89; Stuttford 1994:28). Poor health therefore imposes immense economic costs on individuals, households, and society at large. According to a house survey, the cost of illness in Côte d'Ivoire in 1987 equalled almost 15 percent of per capita GDP, and 11 percent of the workers' normal monthly earnings (World Bank 1994:24). In greater Africa, direct and indirect costs of malaria amounted to US\$800 million in 1987 and were expected to reach US\$1.8 billion annually by 1995 (International Development Research Centre 1996).

1.4 THE CONTRIBUTION OF GEOGRAPHY AND GIS IN STUDYING MALARIA

The impact and magnitude of malaria are affected by a variety of elements. In particular, the interplay of numerous environmental variables influence to a large degree the distribution, abundance, development, activity and longevity of pathogens, zoonotic reservoirs of infections, and their interaction with humans (Center for Health Applications of Aerospace Related Technologies 1997; Martens, Niessen, Rotmans, Jetten & McMichael 1995:458-459). Since

these environmental factors often vary greatly over space, the disease also inherits an important environmental and spatial dimension (Loslier 1995:13). According to Stuttaford (1994:24), “the capability of incorporating a multitude of environmental factors in studying disease patterns ... [would be] a major advance for the [epidemiological] field – especially for vector-borne diseases”. However, although several malaria epidemic models have been developed, yet, “still in their infancy are models evaluating the epidemic hazard to a given area, as based on a set of environmental factors determined for different geographical locations” (Löffler & Malkhazova 1990:360).

Medical geography, the study of the spatial human-environmental interrelationship of disease, nutrition and medical care systems (Barret 1986:27), can make a significant contribution to this challenge, by its main fields of study, namely that of:

- disease ecology, attempting to elucidate the social and environmental causes of ill-health; and
- the geography of medical care, which is concerned with the consumption of care with respect to matters such as distribution and accessibility (Jones & Moon 1987:1-3).

The end product of the geographical approach, or spatial health research, is often attained through its focus on the mapping of diseases, the correlation of spatial distributions by comparing two or more variables (called associative analysis), and spatio-temporal analysis (Douven & Scholten 1995:117).

Despite its importance, the study of environmental determinants of malaria has been hampered by the difficulties related to collecting and analyzing large amounts of environmental data over large areas, and to the speed of change in the malaria epidemiological situation. In this regard the feasibility of studying the environmental determinants of malaria is especially enhanced by Geographic Information Systems (GIS), with its ability to integrate large digital data sets on the basis of their spatial characteristics, and to incorporate data obtained from maps, satellite images, aerial photographs and field work (Brétas 1995:90; Scholten & De Lepper 1991). This spatial information system has the ability to rapidly retrieve, update, query, analyze, model and display diverse data associated with a geographical area. This empowers health managers, researchers and donors through timely and pertinent information, enabling them to make better

decisions, and to plan and prioritize the monitoring, evaluation and combating of the tropical disease (Stuttaford 1994:59). Better estimates of burden of disease, at-risk populations, their distribution and periodicity of annual exposure enables, for example, the effective allocation of limited resources (Medical Research Council 1997; Snow, Marsh & Le Sueur 1996:455-456; Loslier 1995:19; Kitron, Pener, Costin, Orshan, Greenberg & Shalom 1994). However, although being a powerful tool capable of transforming how information is handled, GIS is not a magical solution to all the information difficulties of malaria control. Many analytical techniques are not incorporated within a GIS, and a dual, or linked, spatial analytical system between GIS and statistical modules therefore had to be employed within this study.

1.5 AIMS AND OBJECTIVES

The challenge – searching for malaria epidemiological models and associations as based on a set of environmental factors for different geographical locations – has already been responded to by a research collaboration, entitled “WARDA/WHO-PEEM/IDRC/DANIDA/Government of Norway Consortium Research Project on the Association between Rice Production Systems and Vector-borne Diseases in West Africa”*. The Consortium is currently investigating the health impact of wetland based rice cultivation systems (Figure 1) as a predominant environmental contributory factor to malaria morbidity and transmission, within three different ecological zones in Africa, namely the humid forest lowlands, the Savannah and the Sahel. The overall expected outcome of the Consortium’s activities is “the development and promotion of rice ecosystem management technologies for disease control which make irrigated rice production and cultivation of inland valleys safer and more attractive options for smallholder farmers in West Africa” (Teuscher 1996). In order for effective control interventions to be derived, this study aimed to contribute to the Consortium’s analytical capacity and

* **Consortium members:**

Côte d’Ivoire: Centre Universitaire de Formation en Entomologie Médicale et Vétérinaire, Bouaké.
 Institut Pierre Richet/Organisation de Coördination pour la lutte contre les Grandes Endémies (OCCGE), Bouaké.
 West Africa Rice Development Association (WARDA), Bouaké.

Mali: Faculté de Médecine, de Pharmacie et d’Odonto-Stomatologie, Malaria Research Training Centre, DEAP, Bamako.
 Institut d’Economie Rurale, Niono/Bamako.
 Institut National de Recherche en Santé Publique, Bamako.

Financial Support: Danish International Development Agency (DANIDA), Denmark.
 International Development Research Center (IDRC), Canada.
 Royal Government of Norway.



Figure 1: Wetland rice cultivation system

understanding of the spatial and temporal variations in disease transmission and morbidity, and its spatial interaction with possible environmental determinants. The end result may guide ongoing fieldwork on directions to pursue, result in hypotheses formulation, and assist in the evaluation of the effectiveness of existing malaria control measures.

In order to accomplish this, the study's objectives were as follows:

- To map the spatial distribution of malaria morbidity and transmission risk in a rice producing area of Côte d'Ivoire, thus allowing the rapid detection of the presence or absence of an agro-ecological transmission related risk;
- To analyze the nature of its spatial variation over the study region through spatial data analysis and GIS functionality;
- To investigate the possible influence that a variety of environmental factors might have on malaria morbidity and transmission intensity, and establish predictive models if possible;
- Since epidemiological surveys are very labor and time intensive, the relevant environmental determinants and established models could be used to predict malaria morbidity and transmission in areas with no disease estimates; and

- The results could then be further utilized in an application of the second course of medical geography, namely the investigation of the accessibility and sufficiency of current medical facilities, in view of the predicted malaria related risk.

1.6 AN OVERVIEW OF THE STUDY AREA

In central northern Côte d'Ivoire, the range of rice cultivation systems and ecological settings found are representative of large areas within West Africa where rice cultivation is currently practiced or being promoted. The study consequently focused on a selection of 24 villages situated in the moist savanna of central northern Côte d'Ivoire, in the Korhogo and Katiola districts (Figures 2 to 5). These villages were randomly selected by the WARDA-based Consortium Coordination unit from a sample frame of villages – eight located in each of three 3600 km² agro-ecosystems, characterized by:

- Wetland rice systems where two annual rice harvests are produced, with partial or total water control;
- Rainfed wetland rice systems where a single annual rice harvest is produced, with no water control; and
- Wetlands with no rice cultivation.

Meteorological conditions for the area, on which the flat topography virtually has no influence, are influenced by two distinct (wet and dry) seasons, as summarized in Table 1.

Table 1: Meteorological conditions within the study area

Meteorological Factor	Dry Season (Oct – May)	Wet Season (Jun – Sep)
Temperature Range (°C)	21.0 – 34.2	21.3 – 30.2
Relative Humidity Range (%)	10.0 – 39.1	64.7 – 98.6
Total Rainfall (mm)	360	840

(Source: Teuscher 1999)



Figure 2: Example of a typical village included in the study



Figure 3: Life within a village selected for the study

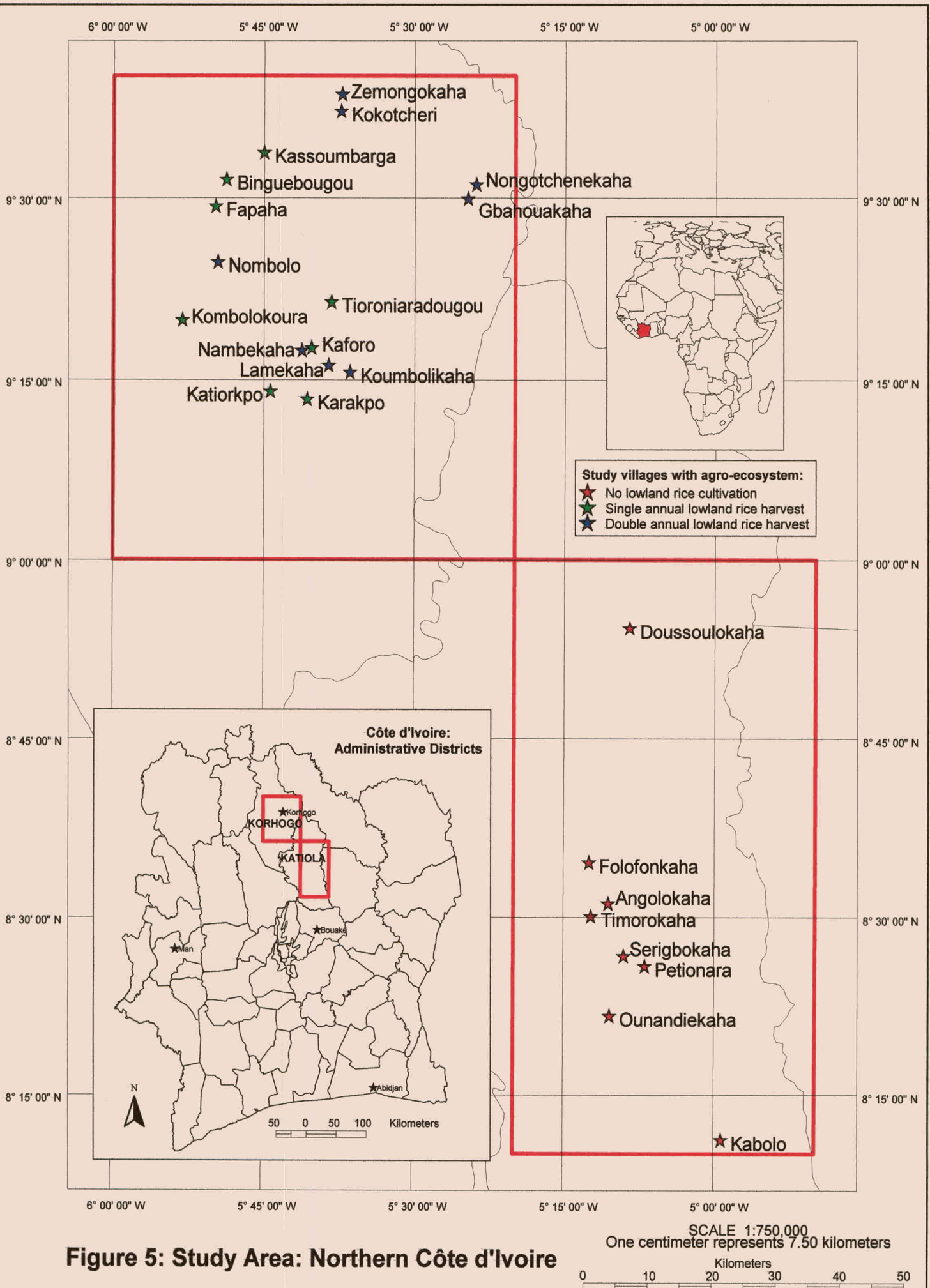
By limiting the surveys to populations within a density zone of 20-30 inhabitants/km², it was hoped to exclude health services and market accessibility as determinants of ill health (Teuscher 1999). With household incomes ranging between US\$160 – 970 per year, inhabitants (from the Senoufo and Tagwana ethnic groups) mostly reside in mud houses (a few in brick). An estimated 50% of the houses has sheet-metal roofing, with the remainder covered in straw thatch. Nearly none perceives a need for mosquito netting (Teuscher 1999).



Figure 4: The moist savanna of central northern Côte d'Ivoire

1.7 CONCLUSION

It is concluded that malaria is indeed still a major health risk, especially in Africa. Malaria research of a spatial nature, which also focuses on environmental factors influencing the malaria cycle, is, however, generally scarce. This study therefore aims to contribute to this information gap through the application of exploratory spatial analysis and spatial functional capabilities of GIS on malaria and environmental data collected within the moist savanna of northern Côte d'Ivoire, in search of predictive models.



CHAPTER 2: MALARIA AT A GLANCE

In order for relevant models and conclusions to be derived, a clear understanding of the malaria cycle, its members (in particular its vectors) and environmental factors which might influence the balance, needs to be attained. This chapter is therefore developed along an elaboration of each of these aspects in turn.

2.1 THE MALARIA CYCLE

Malaria is an infection produced in humans by the *Plasmodium* parasite. As shown in Figure 6, the malaria parasites are transmitted naturally by the bite of a female anopheline mosquito – requiring blood meals to produce fertile eggs. After the mosquito has ingested a blood meal (a) containing the infectious forms of the parasite, the parasite starts the sexual stage of its life-cycle (b), penetrating and creating cysts on the outer wall of the mosquito's stomach (c). These

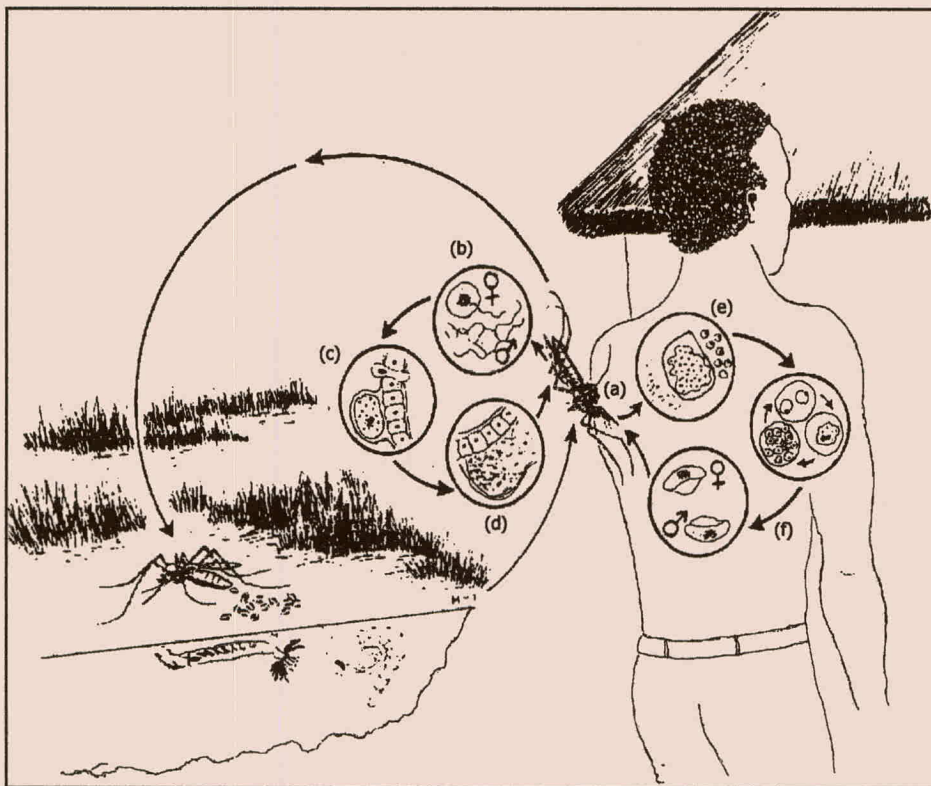


Figure 6: Life-cycle of the Plasmodium malaria parasite in the human and mosquito hosts (Appleton, Sharp & Le Sueur 1995:235)

cysts rupture and the organisms (sporozoites) make their way through the body cavity to the salivary glands (d) where they lie in wait until the infected female bites another susceptible human being (a). The parasite completes the asexual part of its life-cycle in the tissue and red blood cells of humans (e), reproduces (f), and is returned from human to a female mosquito via another bite and blood meal (a), thus completing its development. The symptoms of the malaria attack include alternate bouts of chill, followed by high fever, sweating and prostration, which may last for 5-15 days (Cliff & Haggert 1988:109; Campbell 1998:11). Severe manifestations include cerebral malaria and malarial anaemia, often ensuing in death if untreated.

2.2 THE MALARIA VECTOR

In Africa, the feared carrier (vector) of the *Plasmodium* parasite, the female anopheline mosquito, mainly consist of three mosquito species, namely that of *Anopheles gambiae*, *Anopheles funestus* and *Anopheles arabiensis*. The six subspecies of the *An. gambiae* complex are regarded as being the most efficient vectors of human malaria in sub-Saharan Africa, and often considered the most important in the world. *An. funestus* is also capable of producing very high inoculation rates in a wide range of geographic, seasonal and ecological conditions (American Association for the Advancement of Science 1991:5). *An. gambiae* and *An. funestus* have also proven predominant in transmitting the malaria parasite within the study region (Medical Research Council 1997).

Anopheles gambiae prefers breeding in temporary sunlit pools of cloudy water – closely associated with human activities – and vector populations often increase explosively at the beginning of the rainy season. *An. funestus* are usually associated with more permanent natural sources of water: the edges of rivers, streams, swamps, and the like, usually occurring in clear water with emergent or overhanging vegetation (Gillet 1989:38; Appleton, Sharp & Le Sueur 1995). Adult densities often peak in the middle of the rainy season.

2.3 POSSIBLE FACTORS INFLUENCING THE MALARIA CYCLE

The malaria transmission cycle, and consequently morbidity levels within a region, are influenced by the physical, ecological, agricultural and socio-economic variables of the agro-ecosystem under consideration. Onori & Grab (1980) distinguish between:

- a) **Direct factors**, having a direct influence on the malaria transmission process by affecting any of the three living elements needed for the transmission cycle, namely, the mosquito, the parasite, and humans. These include factors such as:
- the vector density,
 - the human-biting frequency,
 - the parasite rate, and
 - the degree of acquired human immunity to malaria in environments of intense transmission;
- b) **Indirect factors**, creating conditions for an increased or decreased malaria transmission potential, for example:
- meteorological factors, such as rainfall, temperature and relative humidity;
 - vector control (or proliferation) through agricultural practices or environmental management (or the lack thereof), for example alternative wetting and drying of fields, quick flushing, drainage and filling, vegetation control, pesticide use and conservation of natural predators);
 - mosquito resistance to insecticides;
 - parasite resistance to drugs;
 - personal protection (e.g. use of impregnated bed nets, repellents and protective clothing);
 - access to and availability of antimalarial drugs (chemoprophylaxis) and medical treatment (chemotherapy); and
 - housing structures (e.g. use of mosquito screens) (Carter, Brook & Jewsbury 1990; MacCormack 1984; Roger & Bhuiyan 1990:10-14; World Health Organization 1980:7-21).

It is clear, therefore, that environmental criteria as such are by no means the only influencing variables of the temporal and spatial distribution, or intensity, of malaria incidence or

transmission. However, this study's intent is to investigate what proportion, if any, of the disease's distribution might be explained by wetland and rice cultivation related environmental factors. The WARDA Health Research Consortium currently collects a variety of environmental, agricultural, health and social data in the study villages – variables chosen according to literature and relationships anticipated by research members. A brief description of factors addressed in this study follows, starting with a direct factor, namely malaria transmission intensity, followed by numerous indirect factors.

2.3.1 Malaria Transmission Intensity

As part of the malaria cycle, it could be expected that an increase in the transmission level would naturally lead to an increase in malaria occurrence. It is therefore often hoped that interventions which aim at reducing malaria transmission can reduce on a long-term basis the malaria morbidity and mortality. Trape & Rogier (1996) showed that for low levels of transmission (i.e. 0.01 and 0.1 infective bites per person per year) the incidence rate of malaria attacks is probably directly proportional to the level of transmission. However, in endemic areas where malaria is stable, morbidity ceases to be proportional to the level of transmission when transmission varies between 0.1 and 1.0 infective bite per person per year, and could even decrease for very high transmission levels (possibly due to the development of immunity early in life). It was concluded that "...there is therefore no marked variation in malaria mortality according to transmission when this [transmission] is at least equal to one infective bite per person per year" (Trape & Rogier 1996:6; Coosemans & Mouchet 1990; Birley 1991). Several studies in different ecological settings confirmed the lack of a proportional relationship between malaria morbidity and mortality, and transmission intensity, when transmission levels were high (Snow, Bastos de Azevedo, Lowe, Kabiru, Nevill, Mwankusye, Kassiga, Marsh & Teuscher 1994; Snow & Marsh 1995; Snow, Omumbo, Lowe, Molyneux, Obiero, Palmer, Weber, Pinder, Nahlen, Obonyo, Newbold, Gupta & Marsh 1997).

Since malaria morbidity therefore does not necessarily correspond to the level of transmission, the relationships of both malaria morbidity and transmission with possible indirect causal factors are investigated in this study. The indirect factors focused upon are consequently discussed.

2.3.2 Wetland Rice Production's Influence

Being the primary food staple for over 60 percent of the world's population, rice is the most widely grown crop throughout the tropics. About two thirds of the world's 143 million ha of rice land is used for wetland rice culture, in which water from rainfall and irrigation sources is conserved in the field and drainage channels to maintain a shallow flooded condition for almost the entire duration of the crop season (Figure 7) (Roger & Bhuiyan 1990:2). Unfortunately this aquatic environment also provides favourable and extended breeding habitats for several major disease vectors, of which the anopheline mosquito, the vector of human malaria, poses the greatest threat to public health (Wood, Beck, Washino, Hibbard & Salute 1992:2813; Le Sueur, Sharp & Ngxolo 1992).



Figure 7: Rice paddies, serving as extended vector breeding sites

Several accounts where mosquito numbers increased around irrigated regions could be cited in literature, e.g. the Kano Plain in Nigeria, the Ahero irrigation project in Kenya (Stuttaford 1994:27), the Nile delta in Egypt, the Gezira region of Sudan (White 1976, as quoted by Löffler & Malkhazova 1990:356), and the Mamfene area, South Africa (Sharp, Quicke & Jansen 1984; Sharp 1991; Le Sueur, Sharp & Ngxongo 1992). An increased vector population have frequently led to an increased incidence of vector-borne disease – even extensions into areas

where they were not recorded before. This is often attributed to increased transmission levels, extension of the transmission season due to double, triple or even multiple crops, an increase in people-vector contact due to migrant farmers, or an increase in the parasite load which causes clinical symptoms to become more severe (Verhoef & Bos 1992:15-16; Teuscher 1996; Lichtenberg & Getz 1985:292; Dossou-Yovo, Ouattara, Doannio, Riviere, Chauvancy & Meunier 1994). The adverse effect of these schemes on public health has resulted in the World Bank and the African Development Bank taking decisions not to fund irrigation schemes unless an environmental health study has been carried out (Le Sueur, Sharp & Ngxongo 1992; Appleton, Sharp & Le Sueur 1995:238). Indeed, even African rice farmers have tended to stay away from the wetlands and concentrate more on the uplands, due to health hazards owing to various diseases, more difficult working conditions and low average yields under the traditional cultivation methods (World Bank 1994:25). In Africa upland (rainfed) rice cultivation therefore predominates, although not as productive as wetland production, which yields 77% of the world crop (between 2 to 6 tonnes/ha per season) (Service 1989:240).

Although vector-borne diseases are seen as the biggest health hazard in relation to wetland rice cultivation in West Africa, the relationships between them are complex and often highly site-specific. Influencing factors include the developmental stage of the rice plants, water depth and turbidity, soil type, water regimes, double-cropping and the use of insecticides. *Anopheles gambiae* are, for example, found most abundantly in rice fields when the rice is in early growth. As rice grows taller, *Anopheles funestus* may become the more common anopheline, the species previously having been restricted to weedy and shaded irrigation ditches (Service 1989:240-246).

Although wetland rice cultivation is therefore usually seen to be synonymous with increased malaria transmission, this is not always so. Lindsay, Wilkins, Zieler, Dalv, Petrarca & Byass (1991) for example showed that, although previous studies in Gambia have shown that rice cultivation in the dry season is associated with large populations of malaria vectors (Snow 1983), little or no transmission occurred during this period (probably due to climatic stress). Similarly, Robert, Gazin, Boudin, Molez, Ouedraogo & Carnevale (1985) even found that a decrease of malaria endemicity was observed in the irrigated area of the Kou Valley, Burkina

Faso, in spite of a tenfold increase of vector density (probably due to the predominant *An. gambia*'s low survival rates). High vector densities associated with rice fields therefore do not necessarily result in high transmission of malaria, nor is rice growing today associated with the most serious malaria problem areas in the world (Najera 1988).

Bradley (1988) also concluded that in Africa the role of rice fields in increasing malaria intensity is likely to be limited, but is important in prolonging the transmission season – especially in areas with a double cropping season (Carnevale & Robert 1987; Food and Agricultural Organization 1984). In other words, the second crop has favored an increase in vector numbers that would not normally have occurred, since it coincided with the dry season when suitable breeding habitats do not exist. Yet, in Gambia, it again did not coincide with an obvious increase in malaria as might have been expected (Verhoef & Bos 1992:15-16). According to Reuben (1989:144), it is not necessarily irrigation as such, but untidy irrigation that creates problems.

It is therefore important to maintain a balanced view of the overall benefits of irrigation schemes versus health risks, since most have many economic benefits as well as some positive health effects, such as improved nutrition. Only 1.7 million ha out of the 138 million ha of wetlands, suitable for the cultivation of rainfed wetland rice in tropical Africa, are actually tapped for that purpose. Between 10 and 20 million ha of inland wetlands are located in West Africa (Tran, Ofori & Berney 1995:7-8). With appropriate location-specific technologies (see Tran, Ofori & Berney 1995), these can be successfully exploited. The involvement of the health sector and malariologists in the planning stage of water-extension schemes could ensure that possible disadvantages are outweighed by their benefits (Phillips 1990:16; Appleton, Sharp & Le Sueur 1995:243). A proper knowledge of past experience combined with a solid local database can therefore be helpful to avoid or to minimize adverse effects of rice cultivation in the future (Coosemans & Mouchet 1990:9-11).

2.3.3 Other Possible Influencing Factors

In addition to the possible impact rice cultivation may have on the dispersal and magnitude of malaria, other environmentally associated influencing factors addressed in this study are subsequently discussed.

2.3.3.1 Water Bodies and Flooding Pattern

Water reservoirs, necessary for vector breeding, can include the edges of water bodies, wetlands, swamps, paddies and rivers, dependent on the vector species' preference to water depth, turbulence and clarity, as well as vegetation cover (Gunawardena, Muthuwattac, Weerasingha, Rajakaruna, Kumara, Senanayaka, Kotta, Wickremasinghe, Carter & Mendis 1995:104). Even when rainfall is insufficient to sustain other temporary breeding sites, water pool remnants in river beds may continue to fulfil this function (Birley 1991:3-3).

2.3.3.2 Temperature

Temperature acts as a transmission regulatory force. Temperatures above or below threshold temperatures will increase vector mortality and inactivity, and the development of parasites ceases (Martens, Niessen, Rotmans, Jetten & McMichael 1995:460; Birley 1991:3-2). Optimal temperature for mosquito survival is in the range of 20-25°C, with a general decrease in longevity with rising temperature.

2.3.3.3 Relative Humidity

Mosquitoes prefer a relative humidity above 60%, and their longevity increases with increasing relative humidity. If the microclimate humidity, however, is low, it may result in low insect lifespan, and consequently less effective disease transmission (Birley 1991:3-3).

2.3.3.4 Rainfall

Rainfall plays a crucial role in malaria epidemiology since it provides the media for the aquatic stages of the mosquito life cycle (Löffler & Malkhazova 1990:350), whether through small, temporal puddles of rain-water (even cattle hoofprints) (Appleton, Sharp & Le Sueur 1995), or through the artificial regulation of floodwater within water bodies. Rain may prove beneficial to mosquito breeding if moderate, but if excessive it may flush out the mosquito larvae. If an

area has distinct dry and wet seasons, then both insect vector density and disease prevalence are likely to have seasonal patterns. Rainfall may also increase the relative humidity and hence the longevity of the adult mosquito. In a study by Strebel & Thompson (1990:30-32), it was indeed found that significant positive correlations existed between malaria on the one hand, and the rainfall pattern and relative humidity measures on the other. These covariations are, however, subject to time lags between the environmental cause, and the malaria related effect: 2-5 months in the case of the rainfall pattern, and 0-2 months with respect to relative humidity.

2.3.3.5 Insolation

The amount of sunlight influences relative humidity, temperature, rainfall patterns and consequently the availability and sustainability of breeding pools. It has been recorded that this measure tends to vary between the Korhogo and Katiola districts, and is therefore included as having an indirect influence on the vector population. Mosquito preference to the amount of sunlight (see Section 2.2) is, however, regulated more at a micro-level scale by surrounding vegetation.

2.3.3.6 Distance from Village to Water Body

According to Cliff & Haggett (1988:109) and Birley (1991:C-3), the geographical intensity of the mosquito-borne disease is highest where dense human populations live within flying distance of their aquatic breeding grounds. Service (1993:272) notes that the mean flight range recorded of *Anopheles gambiae* males and females was 0.83 and 1.02 kms respectively, with a maximum range of 3.6 km from the release point. Cliff & Haggett (1988:109) recorded that the usual vector flight range is 1.6 km or less from the source of human blood and infection, while Birley (1991:C-3) estimates their feeding to be within 1-2 kms of suitable breeding sites. A radius of two kilometers around study villages therefore seems to sufficiently include major possible breeding sites of mosquitoes active in it. Dispersal may increase during hot weather or with wind, while reduced dispersal with increased distance can be attributed to (a) a 'dilution' factor, due to their numbers being dispersed over an ever increasing area (thus lowering their density) and (b) a 'loss' factor, due to mortality and settling down of the mosquitoes (Service 1993:733).

2.3.3.7 Vegetation

Vegetation might influence malaria transmission in several ways. The forest edge is thought to provide ideal resting places for the mosquito, thus influencing its survival and the level at which malaria transmission is sustained (Gunawardena, Muthuwattac, Weerasingha, Rajakaruna, Kumara, Senanayaka, Kotta, Wickremasinghe, Carter & Mendis 1995:104). It might also form a vegetation barrier to mosquitoes, since “mosquitoes will tend to disperse less in woods than those inhabiting more open terrain” (Service 1993:718). Rain might generate “puddles” between certain plants' leaves suitable as breeding sites, and puddles, lasting longer within the forest than in the open field, might also serve as extended breeding habitats. Hay, Snow & Rogers (1998) found in a study in Kenya that the Normalized Difference Vegetation Index (NDVI), a measure of vegetation biomass, and malaria cases respond similarly, in direction and magnitude, to changes in meteorological conditions. The peak in vegetation activity occurred approximately one month after the rains, and the peak in malaria admissions became evident after a further month. NDVI measures were consequently used as a predictive tool for seasonal clinical malaria.

2.4 CONCLUSION

It has been revealed within the above-mentioned discussion that the malaria cycle, revolving around the transmittal of the *Plasmodium* parasite between human and mosquito, could be influenced and altered by a wide spectrum of factors. A brief description of a set of direct and indirect influencing factors was given, as preparation to the analysis of the covariance between these and malaria occurrence. The actual data collection of these disease and environmental factors within the study area, its manipulation and consequent data reduction methods, are subsequently discussed.

Chapter 3: DATA COLLECTION, MANIPULATION AND REDUCTION

3.1 DATA COLLECTION AND MANIPULATION

Since the disease and environmental attributes under consideration are, like most other natural properties on the earth's surface, conceptually spatially continuous over the study region, their values have been sampled at only a finite number of the infinity of possible locations – in this case, at the study villages. A subset (January to June 1997) of malaria and environmental related data were obtained from continuous surveys in and around the study villages, performed by the WARDA Health Research Consortium (guided by Dr. T. Teuscher) in collaboration with Institut Pierre Richet/OCCGE (directed by Dr. M.C. Henri & Dr. J. Dossou-Yovo), Bouaké. Four malaria transmission surveys were completed during the dry season period: January to June 1997, and two fever episode surveys during March to June 1997. Data which was accessible for this analysis included transmission (entomological inoculation rate) and clinical (fever) data only. Data was selected and manipulated according to geographical criteria or content using Structured Query Language (SQL), geometric operations involving the computation of distance and areas, and statistical operations using MapInfo, IDRISI, Foxpro, SAS and Statistica software, into the following factors as used in subsequent analysis (*factor name denoted in italics*):

3.1.1 Malaria Health Risk Measures

- **Fever episodes** (*Episodes*): Measurements of elevated human body temperature (temperature of at least 37.5°C) are frequently used in the diagnosis of clinical malaria (usually characterized by a space and time specific *Plasmodium* parasite density cut-off value, e.g. 20,000 parasites per microliter blood) in endemic areas, both for clinical management and in epidemiological studies (Smith, Hurt, Teuscher & Tanner 1995:306). Data was acquired during five days' active case detection of new fever cases/1000 people for each village, during which 48 samples were taken. Since fever episode surveys commenced only six weeks after transmission level measurements started, and also at a different time schedule, fever episode data was matched with the nearest corresponding transmission measurement made during the months March to June 1997, in order to make

correlations between them possible. Incidence rates were calculated by the formula: New fever cases/Follow-up days*1000 people.

- ***An. gambiae* transmission intensity (*Ag_trans*):** Denoted here by the number of infective bites/person/month by *An. gambiae* mosquitoes. Measurements involved the appointment of a human “target” within a village hut, who captured any mosquito seeking a blood meal, using a special (straw-like) insect catching apparatus. Captured mosquitoes were classified according to species, and analyzed whether they are carriers of the *Plasmodium* parasite – in which case it would be recorded as an “infective bite”. During the time period, 64 measurements were taken. This indicator of malaria transmission level, also referred to as the entomological inoculation rate (EIR), were calculated by multiplying the mean daily transmission value by a 30.5 day month.
- ***An. funestus* transmission intensity (*Af_trans*):** The number of infective bites/person/month by *An. funestus* mosquitoes, measured and calculated in the same way as *Ag_trans*.

3.1.2 Water Bodies and its Flooding Pattern

- **Wetland rice cultivation (*Supcus*):** The area (ha) under rice cultivation within a two kilometer radius around study villages, as calculated by actual field measurements (using measuring tape to accurately assess shapes and sizes of paddies).
- **Wetland size (*Superfic*):** Morphological area (ha) of wetland within a two kilometer radius from the village (sum of triangulated areas as observed and measured at representative observation points).
- **Flooded area (*Meanfloo*):** Sum of the mean triangulated areas (ha) which were flooded, as observed at the village's observation points during the month. (Flooding is measured once a week.)
- **Number of flooddays (*Floodday*):** Number of days per month when flooding was observed at any of the observation points of a village, multiplied by 7 (since observations were made once a week, representative for the whole week).
- **Percentage of wetland flooded (*Wfloodpe*):** The mean percentage of the village's wetland flooded in the month, as observed at the observation points (thus a proportion of the potentially flooded wetland).

3.1.3 Meteorological Measures

- **Maximum temperature*** (*Tempmax*): The monthly mean of the temperature maximums for each village.
- **Minimum temperature*** (*Tempmin*): The monthly mean of the temperature minimums for each village.
- **Mean temperature*** (*Tempmoy*): The temperature mean of the month for each village.
- **Maximum relative humidity*** (*Humidmax*): The monthly mean of relative humidity maximums for each village.
- **Minimum relative humidity*** (*Humidmin*): The monthly mean of relative humidity minimums for each village.
- **Mean relative humidity*** (*Humidmoy*): The mean relative humidity for each village.
- **Amount of precipitation*** (*Sumrain*): The total amount of rainfall in each village for each month, in mm.
- **Number of raindays*** (*Raindays*): The number of days with rainfall in each village for each month.
- **Insolation*** (*Insulati*): The monthly average of sunshine duration (in hours).

3.1.4 Distance from Village to Water Body

- **Wetland-village distance** (*Nearwetd*): Distance from the center of the village to the middle of the nearest wetland (whether wet or dry), as measured from 1:50 000 topographical maps.
- **Flooded wetland-village distance** (*Vilwetdi*): The mean of the distance from the village to its wetland observation points which were flooded (being representatives of the wetland). If no observation point was flooded, it is presumed that flooding occurred further than 2km, and a value of 2001m was ascribed.

3.1.5 Vegetation Biomass Measure

- **Vegetation*** (*Vegetati*): The vegetation biomass within a two kilometer village radius, as represented by the mean Normalized Difference Vegetation Index (NDVI). The NDVI was calculated with the assistance of Van de Giesen (1997) from reflectance data of LANDSAT

5 TM imagery (1994) (see Figure 8 for examples) by the formula: (near-infrared – red)/(near-infrared + red). In healthy vegetation, the near infra-red reflectance value is high since it is reflected completely, while reflectance in the red band is low since it is absorbed completely - the difference is therefore large (NDVI around 0.4). In comparison, bare soil's NDVI is low (around 0.0). Since similar seasonal rainfall patterns could be observed between 1994 and 1997, it is assumed that the overall amount of vegetation did not change radically in the past three years since the imagery's capture in 1994.

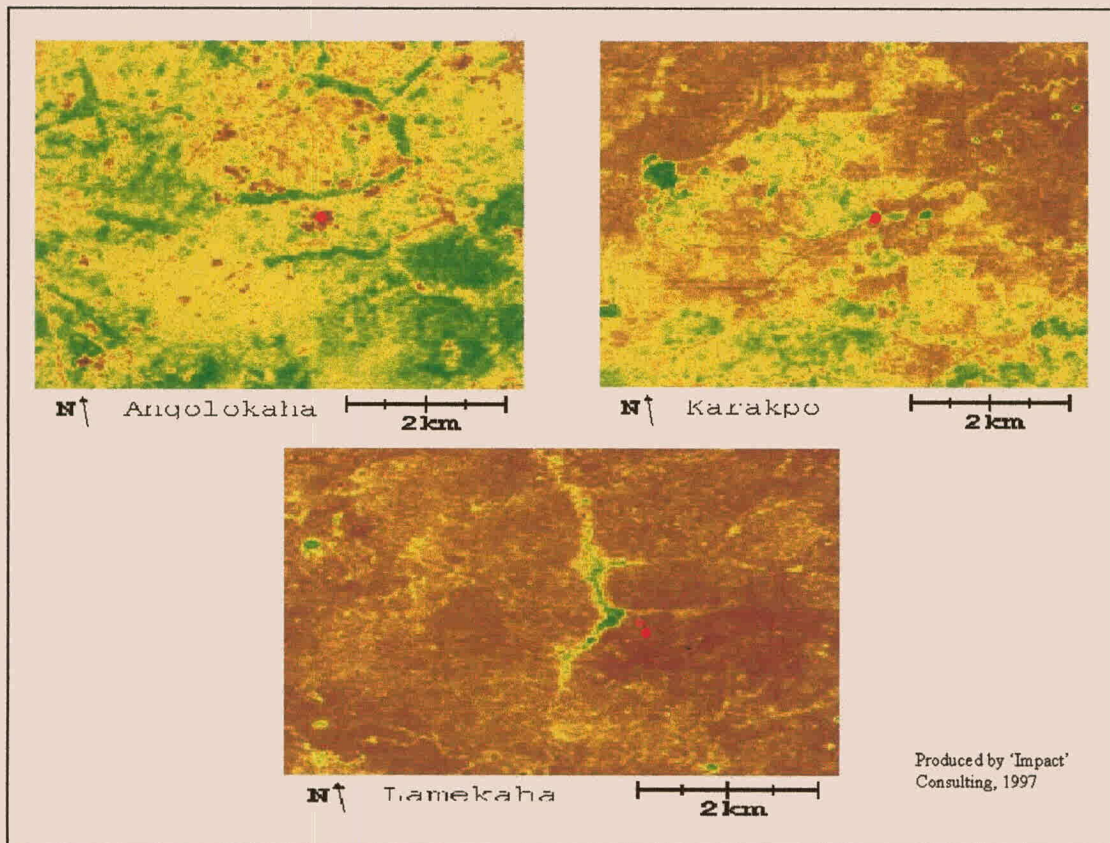


Figure 8: NDVI for some study villages (Van de Giesen 1997) (Villages indicated in red, healthy vegetation (often wetlands) in shades of green, and more barren areas in shades of brown.)

* Notes:

1. Temperature, relative humidity and insolation were measured at weather stations, and were assumed to be the same for all villages within the same district, i.e. for the Korhogo and Katiola districts.
2. Rainfall for Nambekaha was assumed to be the same as the nearby village Kaforo, and for Zemongokaha to be the same as Kokotieri. Rainfall data was interpolated in five instances from the surrounding villages' values, using quintic interpolation in TIN, Arc/Info. Climate within the study region is hardly influenced by the (even) elevation.
3. LANDSAT 5 TM were obtained from Eurimage, Italy: Product Code 984, Path 197, Rows 53 & 54, acquired on 28-04-94 and 12-04-94 respectively.

Since the relative quantitative scales at which the different attributes are measured at a location could significantly affect the results and interpretation of multivariate techniques, the independent variables were standardised before use, so that all variables are measured on the same standard quantitative scale. It was further postulated by the Consortium that the agro-ecosystem surrounding villages, i.e. whether single, double or no rice cropping is cultivated, might have an influence on the disease incidence and malaria transmission. However, by applying one-way Anova to the fever episode data variance, using the agro-ecosystem as a grouping factor, it was found that an F-ratio of 0.0374 would occur more frequently than 96 times in a hundred random samples. The null hypothesis (i.e. that the fever episode means of the three groups are equal in the parent population) was therefore accepted. The fever episode data could therefore not be divided into three “agro-ecosystem” groups in further analysis, but was considered in its totality. The same test, applied to the transmission level of *Anopheles funestus*, resulted in an F-ratio of 5.1124 and an F-probability of 0.0096. The alternative hypothesis, that at least one group’s mean transmission level is different from the others, could be accepted. In fact, as will be seen in Figure 10 shortly, it is suspected that *Anopheles funestus* transmission is primarily concentrated in the villages with no wetland rice cultivation within its vicinity. No conclusion could be made with certainty in the application of the test to the *Anopheles gambiae* transmission level (F-ratio = 0.3084 and F-probability = 0.7361). For simplicity, none of the transmission data was divided into “agro-ecosystem” groups either, but further analyzed in its entirety.

3.2 DATA REDUCTION

A useful way to view the group of explanatory variables, p , having measurements at each of n observation sites, is as a set of n points in a p -dimensional space – also referred to as the attribute space. This needs to be differentiated from the geographical space – the real world spatial location of the observations. The interest during the analysis of these multivariate measurements lies in studying the pattern in these values taken together, rather than in each taken in isolation. Since there are so many independent variables, of which many highly correlated, a data reduction method called Principal Components Analysis was performed. It implies a linear transformation of the original variables to a new, reduced number of variables by rotating the original coordinate axes of the attribute space orthogonally, while making the

variance of the transformed observations (principal components) as large as possible. In other words, the original coordinate system (y_1, \dots, y_p) , was rotated to a new one, (u_1, \dots, u_p) , where u_1 is measured along the direction in which the observations have the most variability, u_2 is measured along the direction that maximizes the remaining variability, and so on. In this manner, consecutive factors are extracted. Since each consecutive factor is defined to maximize the variability that is not captured by the preceding factor, consecutive factors are independent of each other – thus uncorrelated or orthogonal. The relationship between the principal component and the original variable can be evaluated through the covariance between them – often referred to as the loading of the original variable on the principal component. Using these loadings it is possible to interpret what each of the principal components is measuring in terms of the original variables. Applying this statistical transformation to the 17 standardised independent variables (using Statistica Version 5, with a normalized Varimax rotation), the first five principal components (chosen according to the scree test) account for 86.34 per cent of the total variation, with the corresponding variances in each of the new axes accounting for 34.5%, 25.08%, 13.15%, 8.68% and 4.93% respectively (see Appendix A for details). The following variables have a high positive loading on the first component: *Humidmin*, *Humidmoy*, *Humidmax*, *Sumrain* and *Raindays*; and the variable with a high negative loading: *Tempmax*. This component may therefore reflect the main characteristics of the amount of **humidity** in the air. The variables having a high positive loading on the second factor are: *Supcus*, *Superfic*, *Meanfloo*, *Floodday* and *Wfloodpe*. This factor may therefore describe the **size of the water bodies** around the villages. Variables loading positively high on factor 3 are: *Tempmoy* and *Tempmin*. This reflects the **temperature** fluctuations. Variables having a high positive loading on factor 4 are *Nearwetd* and *Vilwetdi*, thus being an indication of the **distance to water bodies**. *Vegetati* have a high negative loading on factor 5, and may differentiate **barren** areas from those with a high volume of biomass. Since each component is by definition the “best” linear summary of variance left in the data after the previous components are accounted for, only the factor score “observations” of the first five new variables are retained for further analysis, knowing that not much information (variance) is lost by doing so (Kim & Mueller 1978; Mulaik 1972).

CHAPTER 4: EXPLORATORY SPATIAL DATA ANALYSIS

Since the focus of this study lies on the exploration of patterns and relations of spatial objects, and not on the establishment and testing of a pre-conceived hypothesis, it was decided to employ an analysis technique described by Douven & Scholten (1995:117, 130) as exploratory spatial data analysis. The exploratory spatial analysis techniques might indeed lead to new insights of the spatial distribution and relations of data and even to the formulation of a hypothesis, to be tested by confirmatory spatial analysis. Spatio-temporal analysis, having a more model-driven character, could be useful in this type of analysis. However, the main focus of this study's exploratory approach and also the progression of this discussion, would be along the recurrent stages identified by Mayer (1986:76-79), Bailey & Gatrell (1995:21-25) and Douven & Scholten (1995:119) within aggregate spatial health research, namely:

- (a) Mapping of data to identify spatial disease patterns;
- (b) Exploring spatial disease variation;
- (c) Describing and measuring the association between disease and other spatially varying factors;
- (d) Modelling or predicting disease; and
- (e) Interpretation of the results, identification of possible areas of further research and generation of possible hypotheses.

4.1 MAPPING MALARIA RISK

After the collection and preparation of health and environmental data, the first stage in spatial health research is to map the data – in accordance with the first objective of this study. The sequential stages identified by Douven & Scholten (1995:123-128) and Mayer (1986:76-79) herein are:

- (a) The delineation of its geographical distribution (cartography); and
- (b) The production of probability maps.

Each of these are subsequently described and applied to the data set.

4.1.1 Geographical Distribution of Malaria

Map-based presentation of human disease has a long history. Disease atlases exist for the rarest of cancers in the developed world, and yet such a basic epidemiological tool does not exist for the single largest cause of mortality in Africa. It is notable that attempts to define the burden of malaria in Africa can be little more than educated guesses, given the lack of specific details on transmission and population distribution across the continent (Snow, Marsh & Le Sueur 1996:455-456). Mapping Malaria Risk in Africa (MARA) is a collaborative effort in addressing this information gap, currently developing a dynamic atlas of malaria risk and burden, down to the second administrative geographic unit (provincial) in Africa (Medical Research Council 1996:1-2; Stadler 1998:20-24; Anonymous 1998:6-9; Campbell 1998).

Malaria risk maps can be informative, intuitive and easier for specialists from all disciplines to comprehend. Its visual impressions often culminate in a new understanding and interpretation of the spatial variation of vector or disease incidence, which may result in hypothesis formulation regarding its covariation with environmental determinants and the health care system (Yoon 1995:202-205; Loslier 1995:18; Kitron, Pener, Costin, Orshan, Greenberg & Shalom 1994). Mapping the geographic occurrence of the disease and its severity could further provide a critical component to the rationalization and communication of appropriate and cost-effective intervention strategies, and the allocation of limited human and financial resources – ideally targeting high-risk areas of need where they are likely to be most effective (Medical Research Council 1997).

Since this study dealt with spatially continuous data sampled at village “point locations”, proportional circle maps of fever episodes and malaria transmission could be created (Figures 9-10). The particular advantage of using proportional circles in mapping distributions is that they can be combined with pie-charts, allowing information about the relative magnitude of different categories to be perceived very quickly (Douven & Scholten 1995:123-124).

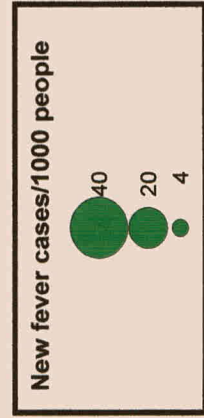
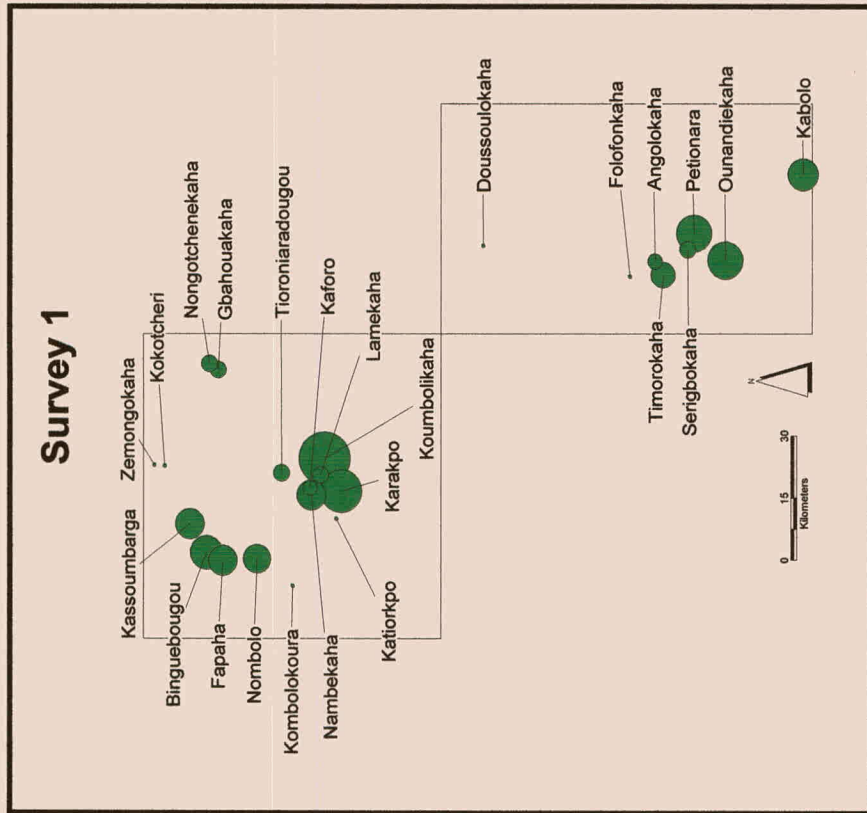
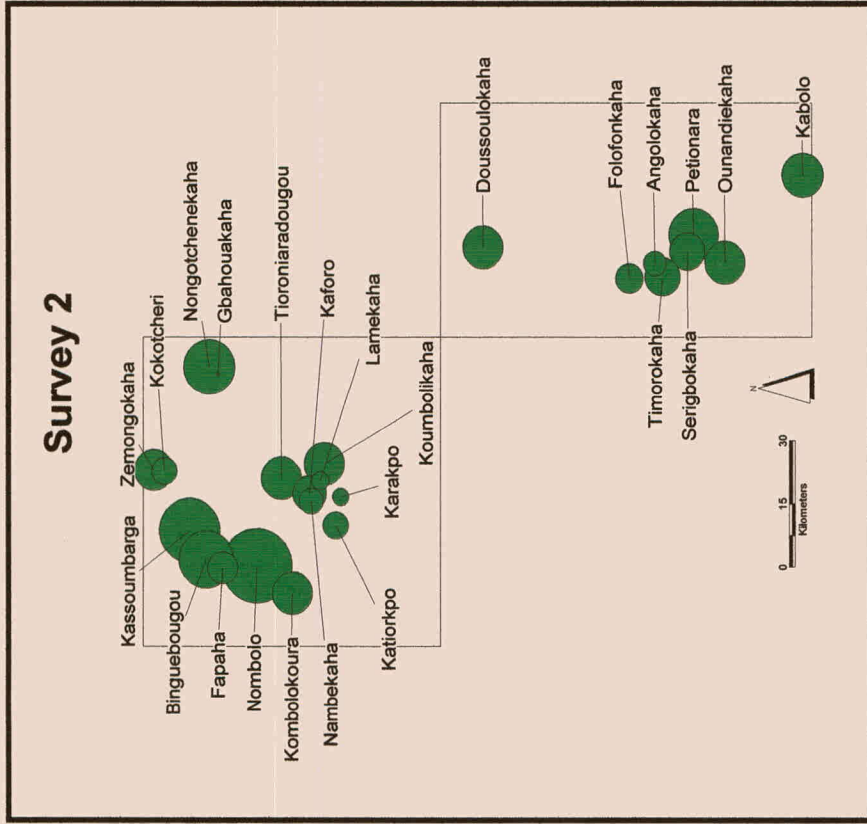


Figure 9
Fever Episodes: Mar - Jun '97

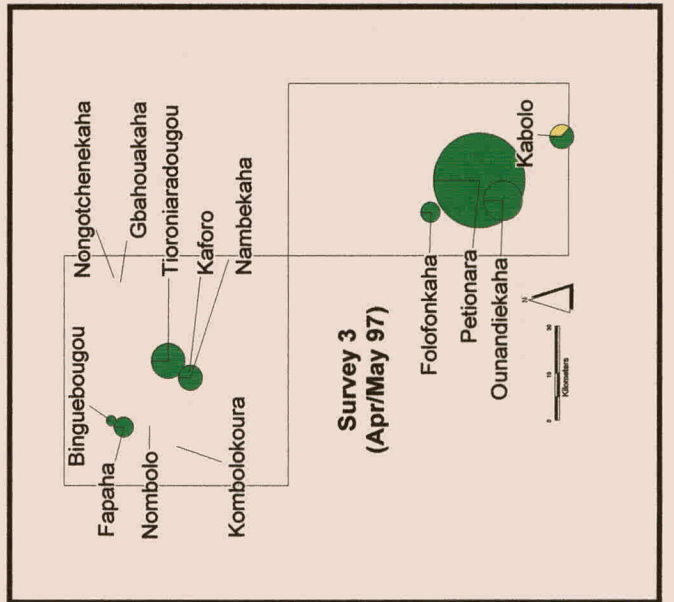
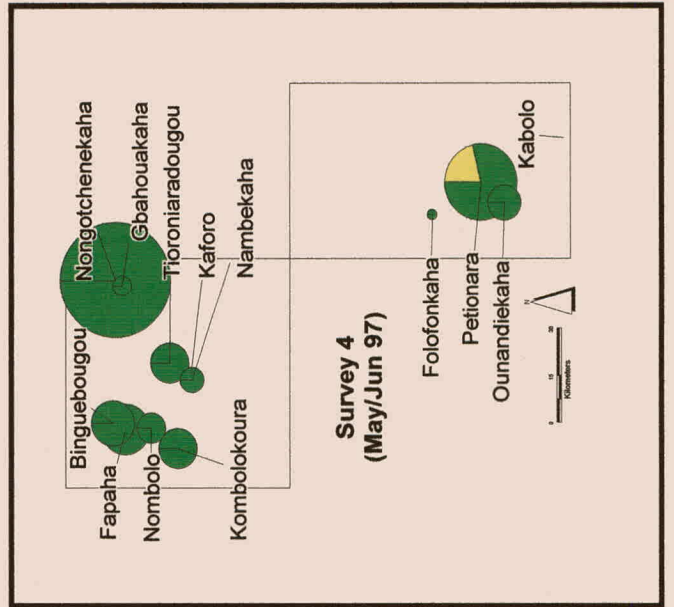
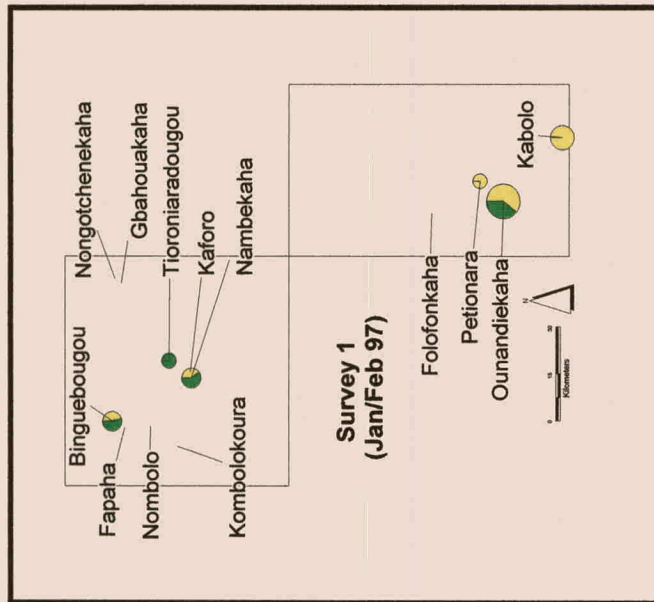
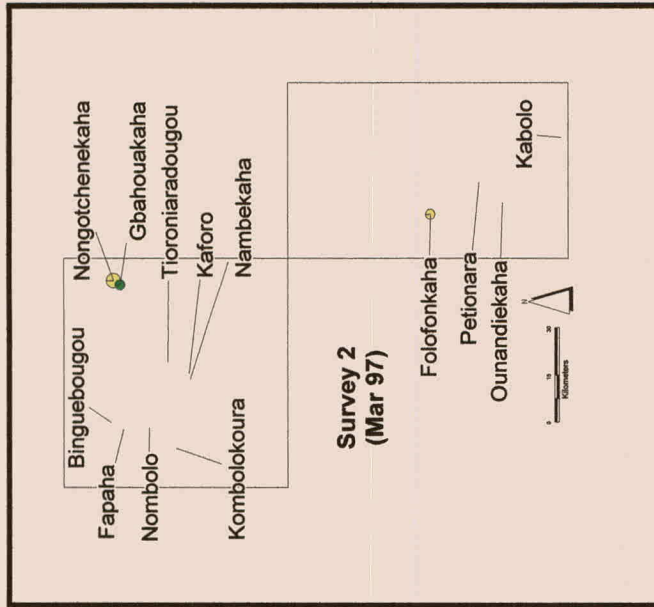


Figure 10
Malaria transmission intensity: Jan - Jun '97

In Figure 9 a general increase in fever episodes could be detected as time progressed between the two surveys. Similarly, although overall transmission levels were low during the first two surveys, it rapidly increased in the third and culminated in a climax in the fourth survey (May/June '97) – the beginning of the rainy season (Figure 10). Since surveys were done in the dry season, this observation stands in agreement with information already gained in Section 2.2: *An. gambiae* populations often increase explosively at the beginning of the rainy season, while *An. funestus* density climaxes only occur in the middle of the rainy season. It also comes as no surprise that *An. funestus* transmission is concentrated in villages without wetland rice cultivation, since their breeding preference is in more permanent natural sources of water, in comparison with *An. gambiae*'s preference to be in close proximity with human activity. The presence of malaria incidence during March/April, even during an absolute minimum transmission level, might confirm Trape & Rogier's (1996) conclusion that disease occurrence cease being proportional to transmission levels when transmission is at least equal to one infective bite per person per year. The concurrent increase in both transmission and fever episodes, however, lead to an anticipation that a positive correlation between disease incidence and transmission intensity exists, predominantly inflicted by *An. gambiae* during this time period. These preliminary observations are subsequently investigated.

4.1.2 Probability Maps of Disease Incidence and Transmission Levels

After the geographic distribution of a disease is mapped, further insights could be attained by determining the areas with extreme data values. It is possible to define respectively significantly high and significantly low areas of disease incidence, or transmission level, using various techniques, as described by Douven & Scholten (1995:125). The method employed in this study involved the determination of data values exceeding the 90th percentile of the normal probability distribution (therefore values exceeding the measure below which 90% of cases are to be found, assuming a normal distribution approximation). The constructed frequency distribution is consequently compared with a normal distribution (Figures 11, 12 and 13), and the fever episode and transmission values determined, corresponding to $P(z \leq 1.28) = .10$. Values above these critical levels could therefore be considered as significantly high (extreme) values. In Figures 14 and 15, showing the result, it is noteworthy that the majority of significantly high fever episodes are concentrated in the second survey, forming a "hot spot" in

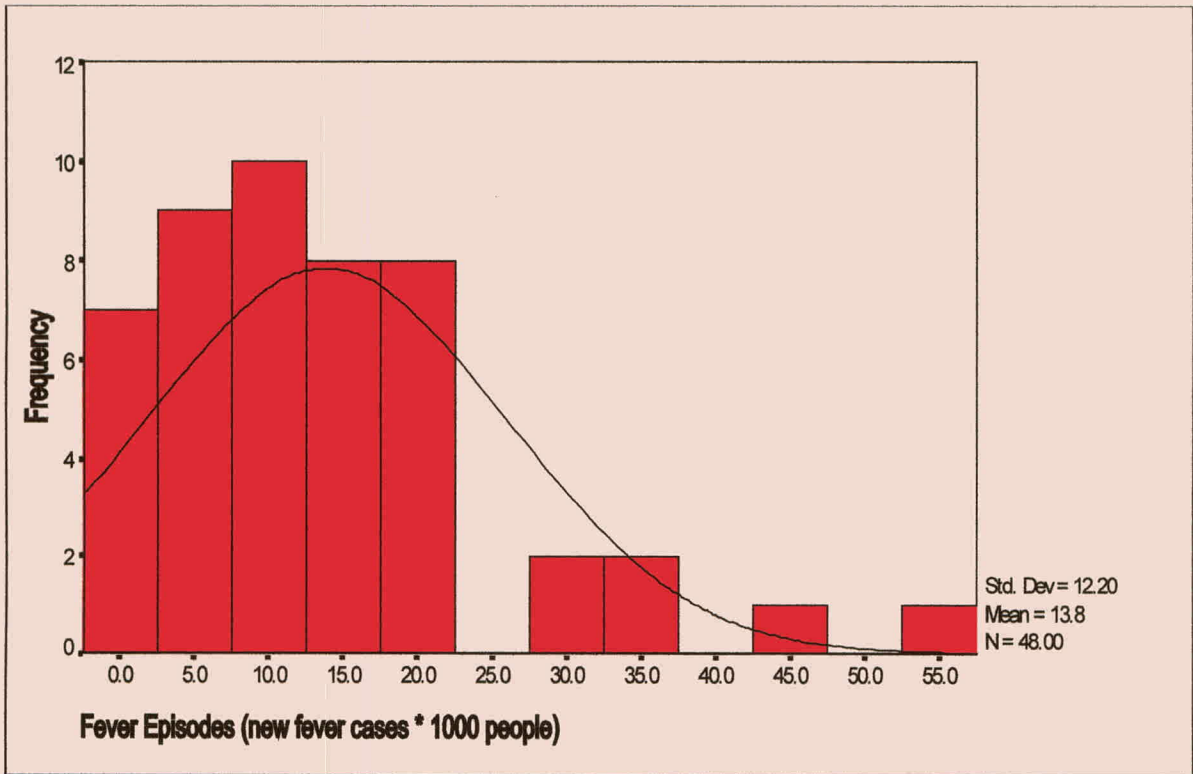


Figure 11: Frequency and normal distribution of fever episodes

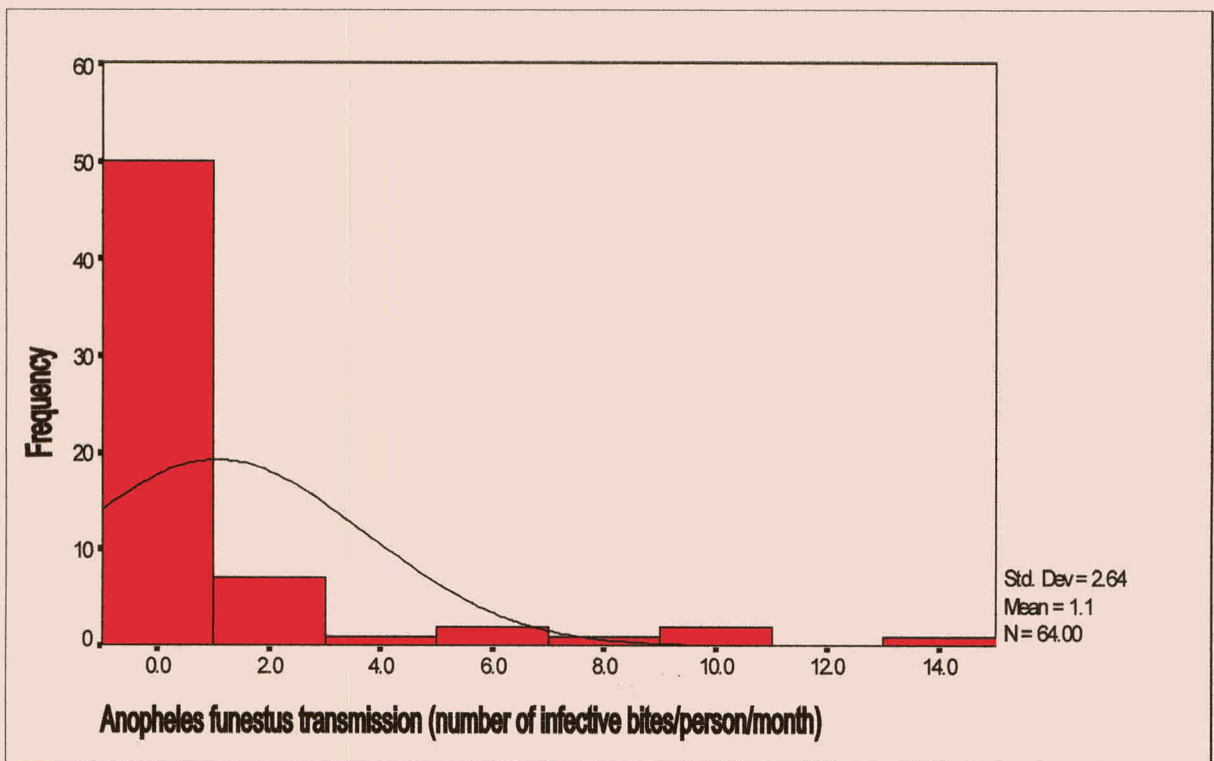


Figure 12: Frequency and normal distribution of *Anopheles funestus* transmission

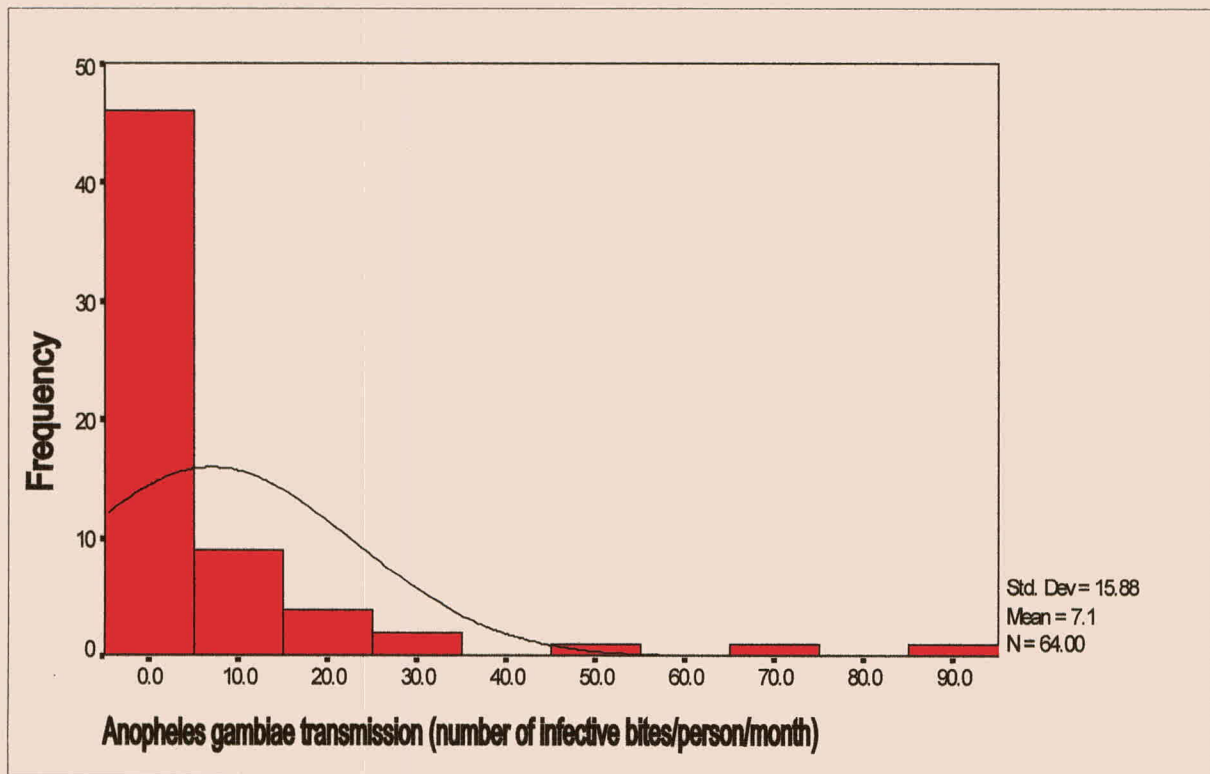


Figure 13: Frequency and normal distribution of *Anopheles gambiae* transmission

the north west of the study region. It is also clear that Petionara and Nongotchenekaha are especially prone to high transmission levels at certain time periods. Areas of high risk (and concern) have therefore been identified, enabling more focused searches for determining factors and the establishment of cost-effective intervention strategies. These extreme cases are further addressed in associative analysis.

4.2 EXPLORING SPATIAL DISEASE VARIATION

Spatial variation in health related data is well known, and its study is a fundamental aspect of epidemiology (Nobre & Carvalho 1995:22). In order to further understand and objectively characterize the spatial arrangement and distributional patterns identified in the disease maps, numerous techniques, as described by Bailey & Gatrell (1995), can be employed. The main objective of the analyses would be to infer the nature of spatial variation in a particular attribute over the whole of the study region, from the sampled point (village) observations. Description in terms of a continuous surface which captures large scale global trends (first order variation in the mean value of the process), as well as the aspects of small scale local variability (second order variation or spatial dependence between observations for any two locations) are sought.

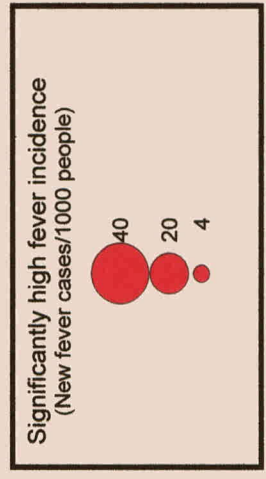
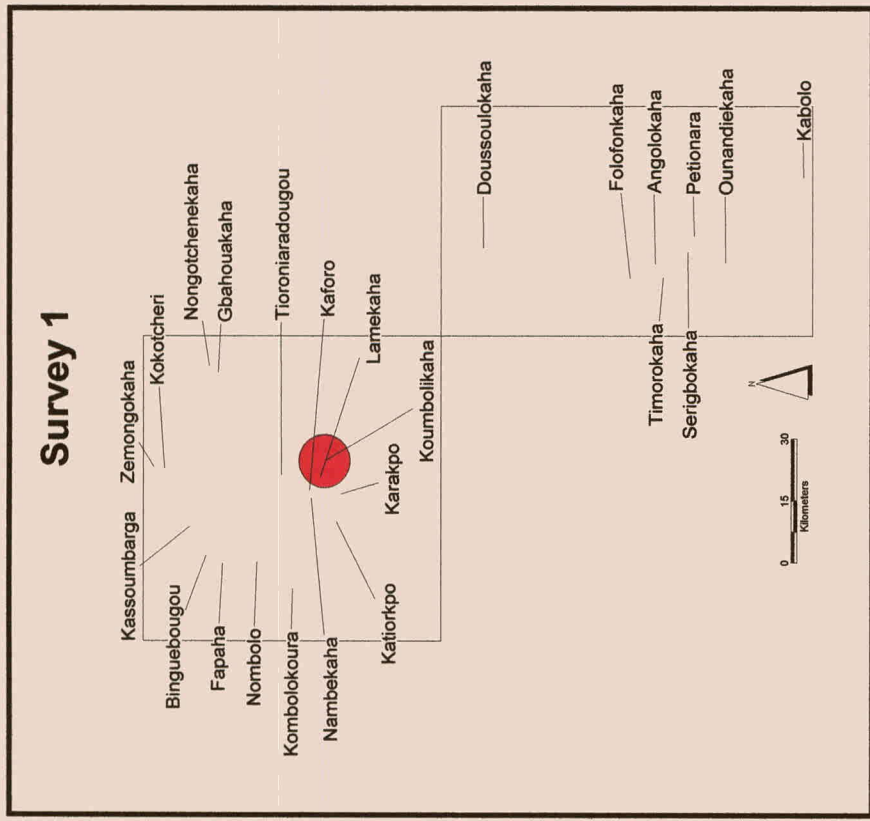
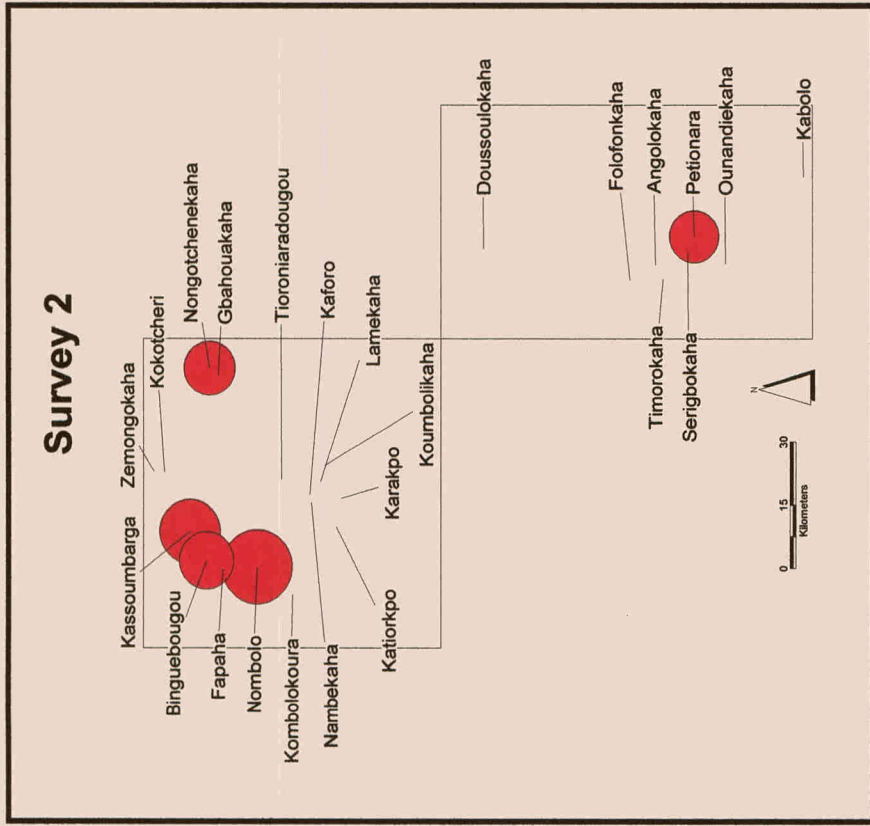


Figure 14
Extreme fever episode values:
Mar - Jun '97
 (Values exceeding the 90th percentile
 of a normal probability distribution)

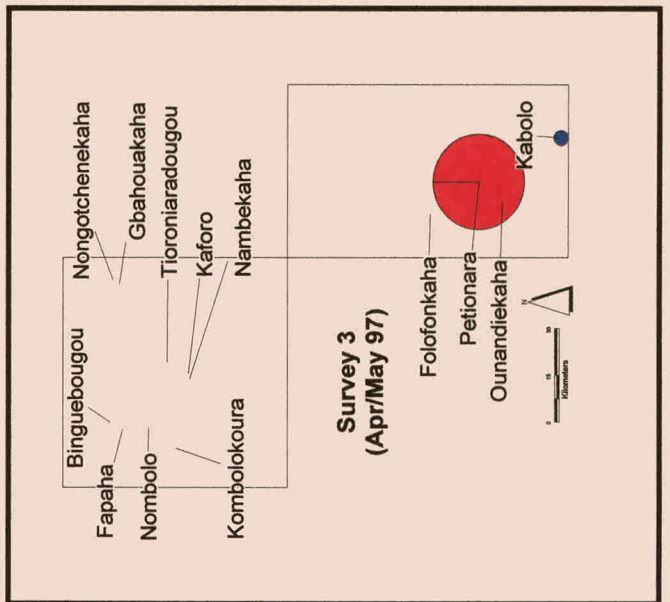
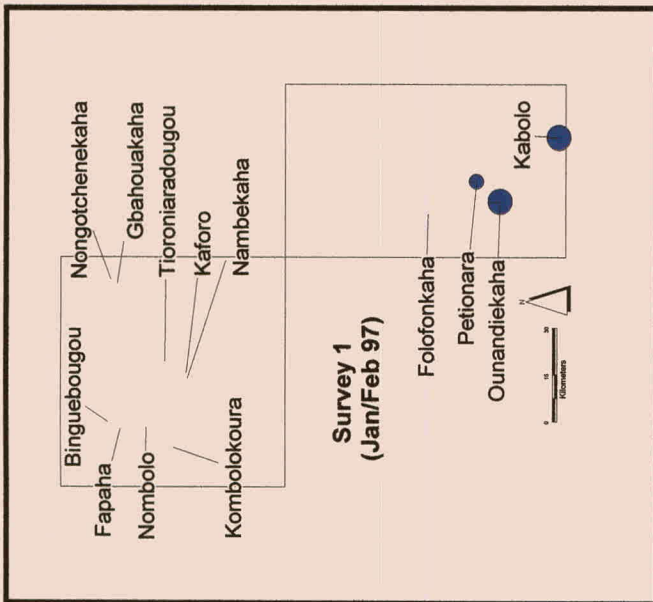
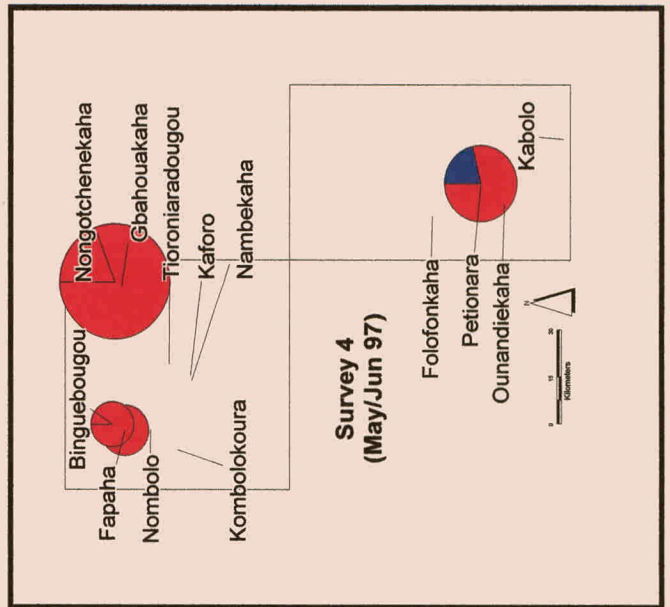
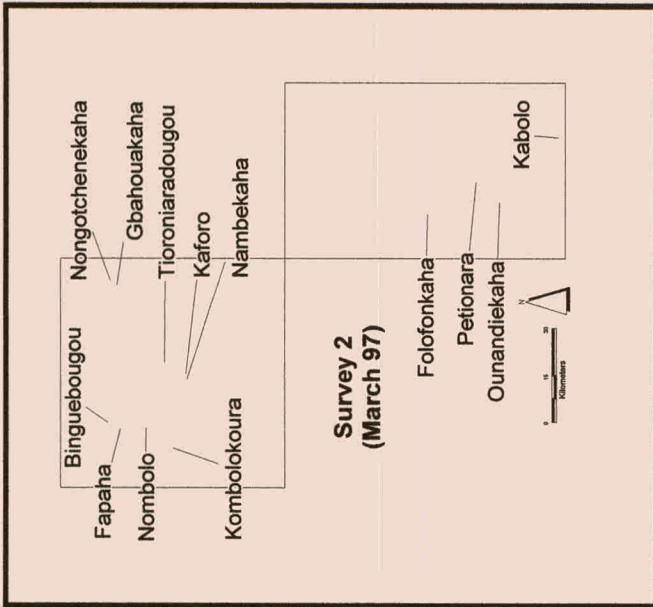
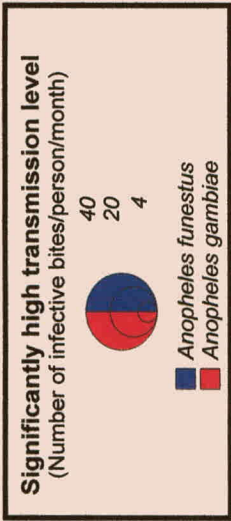


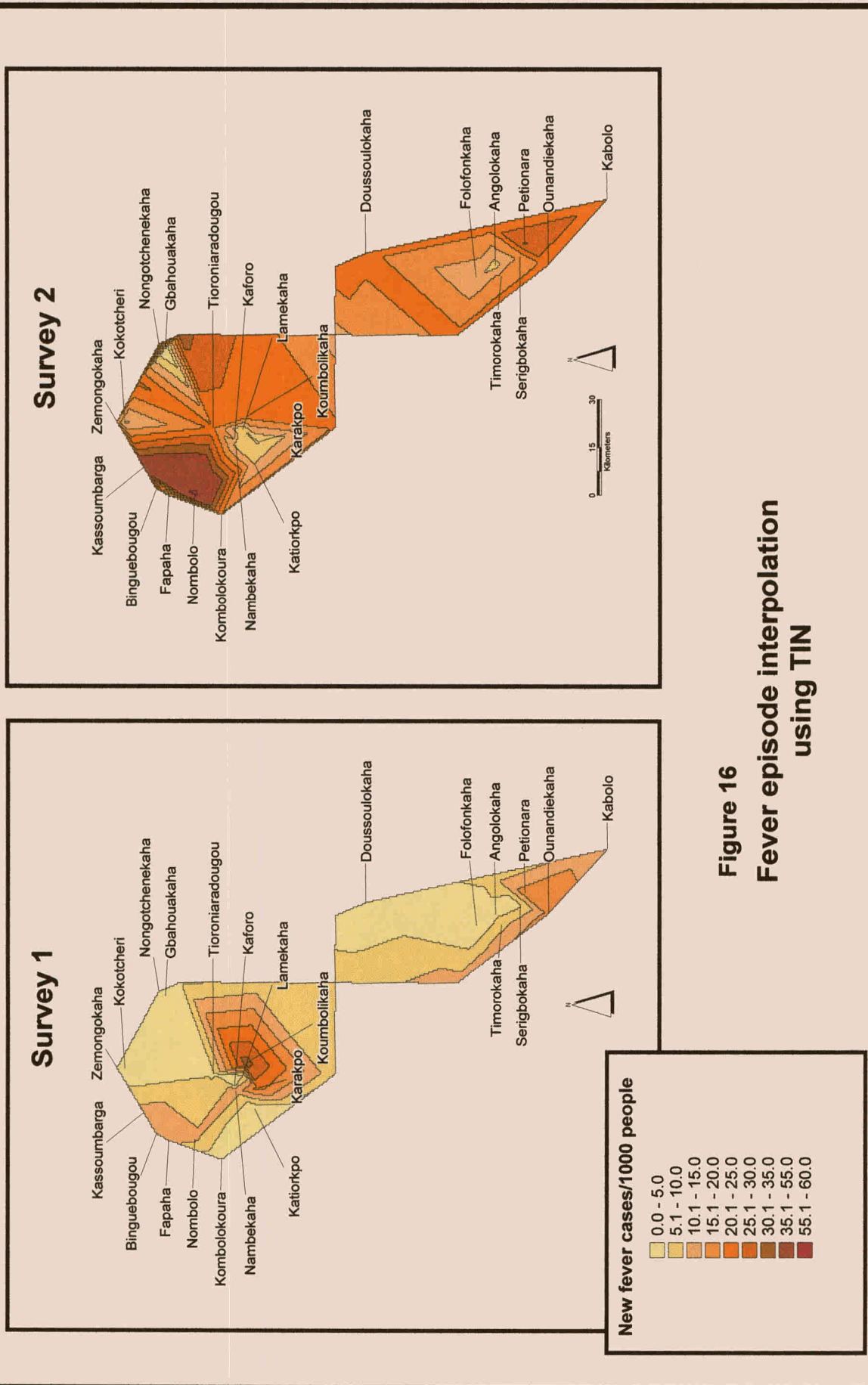
Figure 15
Extreme malaria transmission levels: Jan - Jun '97
(Values exceeding the 90th percentile of a normal probability distribution)

4.2.1 Exploring First Order Variation

In order to investigate how the mean value of the dependent variables vary across the study region (first order variation), the irregularly distributed sampling points in the study area were used to create a Triangulated Irregular Network (TIN) (Environmental Systems Research Institute 1992). In principle, values could consequently be interpolated at any location within the study “convex hull”, and regions of the same level of fever episode occurrence or transmission intensity could be produced. This makes the examination of the variation in the mean value simple (Figures 16-18). The concentration of high fever episode levels in the north west and periodic high transmission risk in Nongotchenekaha and Petionara were again apparent.

4.2.2 Exploring Second Order Variation

Intending to explore the way in which the deviations of observations from their mean value at different point locations co-vary or are correlated (the second order properties), the semi-variogram came under study. Since many spatially continuous phenomena show “spatial persistence”, larger positive covariances or correlations are typically observed between two observations at short distances from each other, than those at greater distances. The variance of the differences between pairs of observations (the semi-variance) is therefore dependent on the distances between the observation points. The graphical plot that relates the semi-variance to the distance is known as the semi-variogram. It is based on the regionalized variable theory that assumes that the behaviour of the observed spatial process in some parts of the space is the same as that in others (thus stationary). Isotropy also had to be assumed in producing the isotropic semi-variograms using the Kriging command of Arc/Info (Environmental Systems Research Institute 1992: 5-5). The fever episode semi-variograms, shown in Appendix B (Figures B1 – B2), reveal a high degree of variability in the data, even over short distances. This may imply an absence of spatial continuity. Within some transmission semi-variograms (Figures B3 – B10), the same conclusion could be made. Within others the semi-variance is horizontal relating to the distance, signifying no spatial dependence between observations. Since none of the semi-variograms rise to an upper bound (a “sill”) beyond which the variation in observations is no longer correlated, some degree of non-stationarity may be suspected in the process. No clear second order variation could therefore be identified, and there would be little



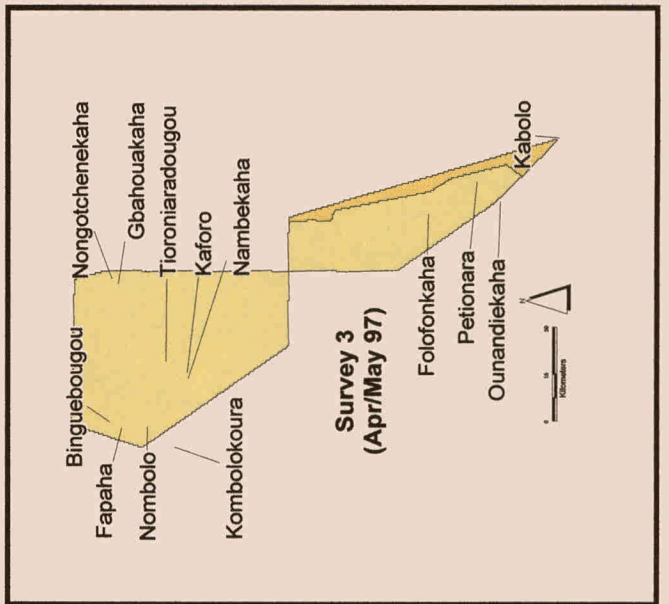
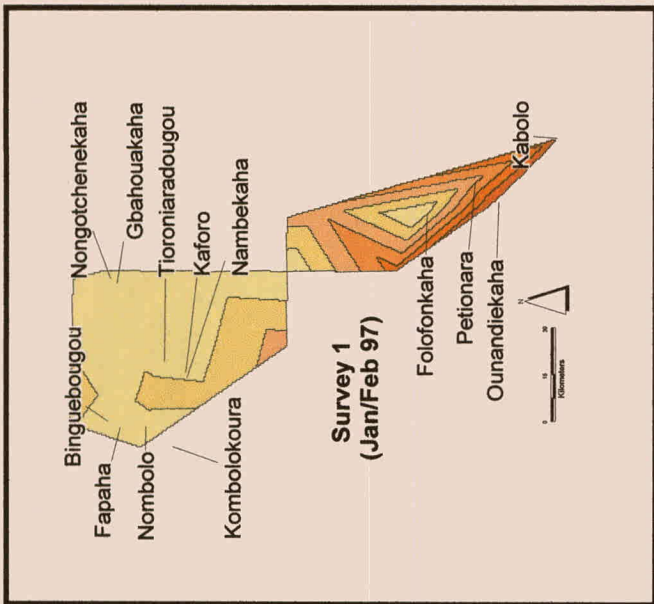
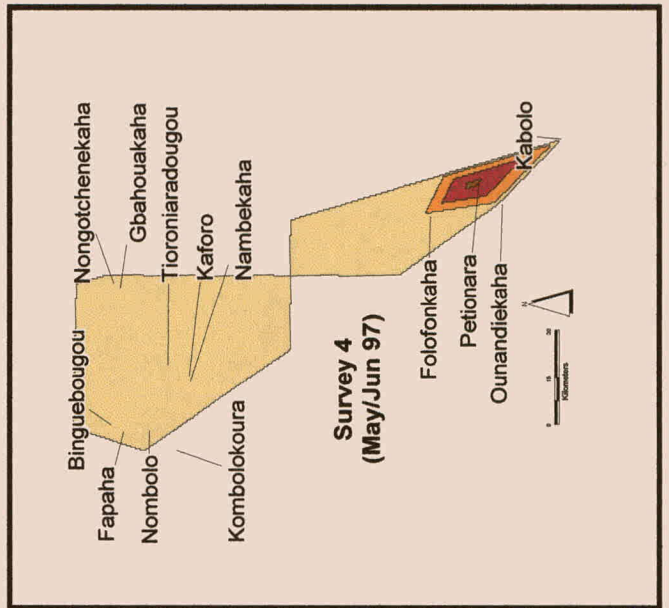
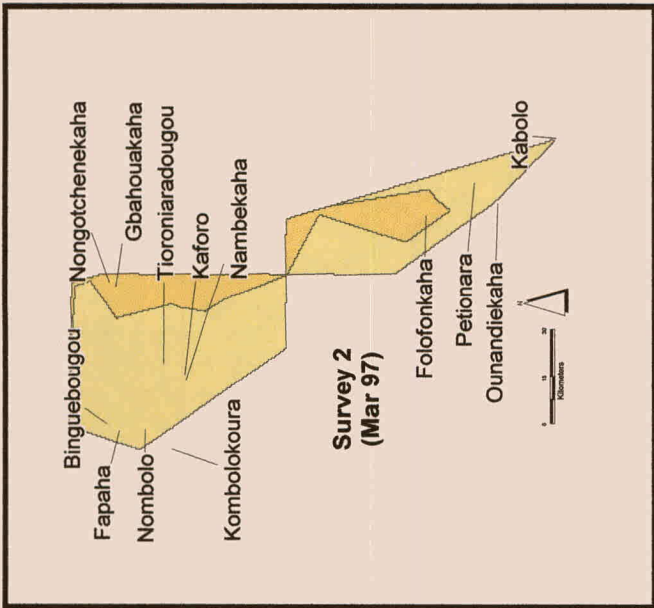
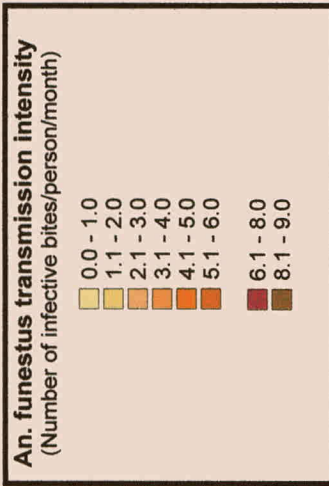


Figure 17
Anopheles funestus
transmission interpolation
using TIN

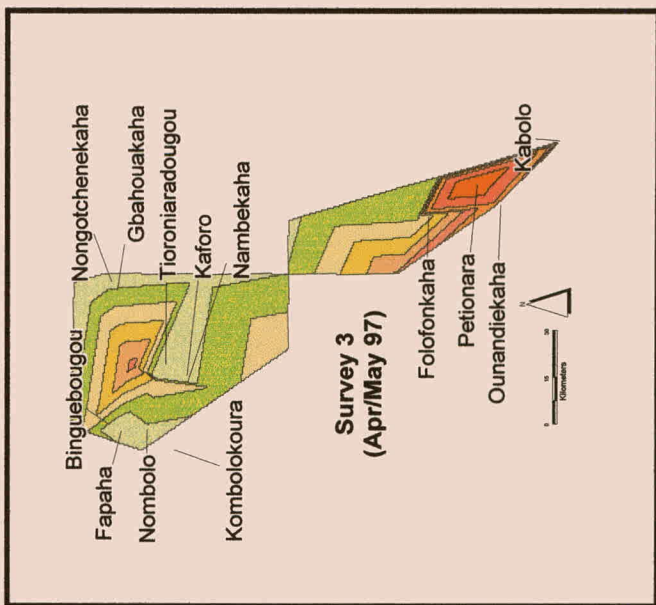
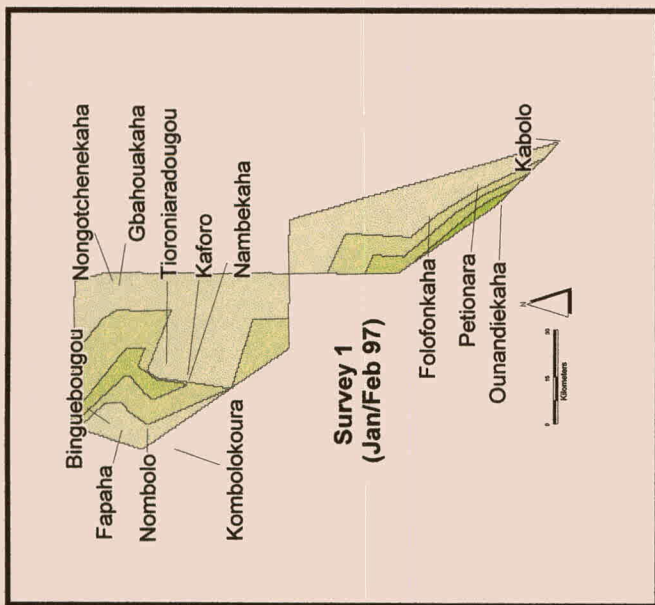
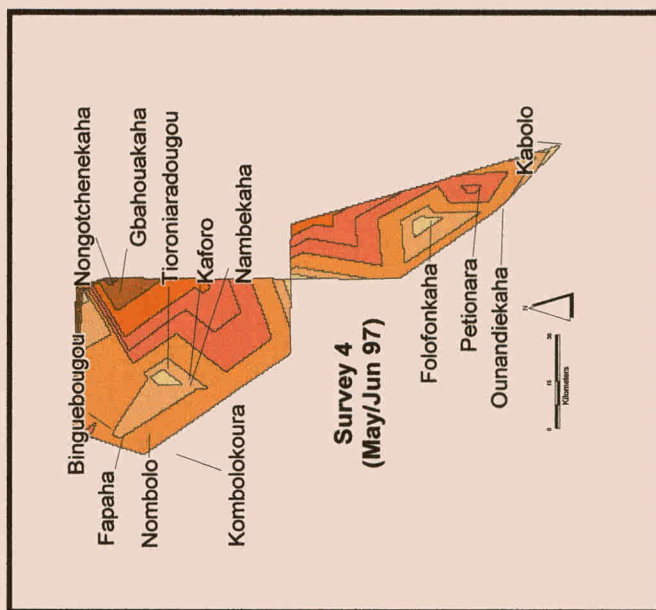
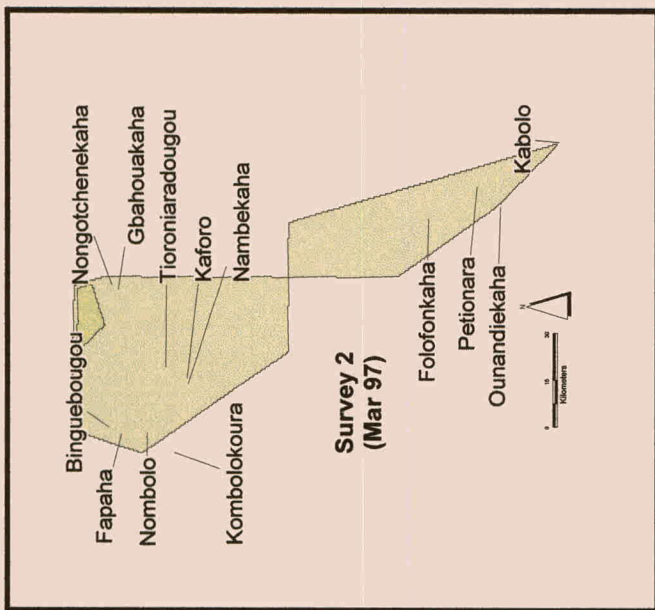
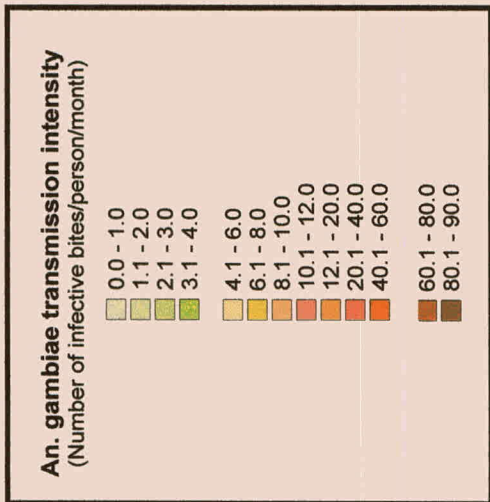


Figure 18
Anopheles gambiae
transmission interpolation
using TIN

point in applying local spatial prediction techniques, e.g. Ordinary Kriging, which include both first and second order effects in the interpolation process. Noting the general global trend would probably be as far as analysis can go in order to understand the variables' spatial distribution.

4.2.3 Modelling First Order Variation

A simple and widely used approach to model global trends (first order variations) is trend surface analysis. It involves the calculation of a single function which describes (or “fits”) the disease related “surface” that covers the entire study area. The dependent variable is therefore approximated by a polynomial expansion of the geographic coordinates of the point (village) observations, with the coefficients β_i of the polynomial function being found through standard least squares regression results. The standard regression assumption is that the random error component ϵ is normally distributed with a mean of 0 and a constant variance σ^2 . Furthermore, it is assumed that the errors are independent, so that their covariance is zero. This implies that there are no second order effects present in the dependent process Y – an assumption that seems to be true given the accompanying semi-variograms. Since there are some intercorrelation between the X and Y coordinates, forward stepwise regression was applied which takes particular note of the problems of multicollinearity, while retaining only those terms with t values that are significant at the specified α level. The resulting regression models were as follows:

Episodes: None of the variables contributed significantly enough to the regression at an entry level F of 1.0. No equation could therefore be deduced.

$$Af_trans = 14.16317 - 0.00001Y_Coordinate$$

$$Ag_trans = 110.7185 + 0.0000723X_Coordinate$$

The accompanying statistics were:

	N	R ² (%)	Adj. R ² (%)	F-ratio	p	df	σ	Durbin-Watson
<i>Af_trans</i>	52	17.68	16.03	10.739	.00191	1,50	1.6043	1.6908
<i>Ag_trans</i>	52	2.07	0.11	1.058	.30867	1,50	16.313	1.3358

R-square indicates the proportion of the variance in the dependent variable accounted for by the regression model. It is implied that almost 18% of *An. funestus* transmission are explained by the model, and 2% of *An. gambiae* transmission. The significance of this can be tested using the **F-test**, where the null hypothesis of no explanation ($\beta_1 = \beta_2 = 0$) can be rejected if the F-ratio is larger than the critical F, or if the significance level for F, namely p, is less than the specified α . At an α of 0.05, the null hypothesis is, however, accepted for the *Ag_trans* regression model.

When the residuals are checked for autocorrelation, it ensures that the data conform to our purely statistical needs. The **Durbin-Watson** statistics indicate whether there is in fact a non-significant autocorrelation between the residuals. If a significant autocorrelation does exist, the regression model cannot be accepted. It is clear from the Durbin-Watson values that the *Ag_trans* model can therefore not be accepted, as the value of 1.33 is not close enough to the non-significant autocorrelation indicator of 2.0.

No higher order polynomial regression could significantly improve the models above, and the only model of any significance is therefore that of *An. funestus* transmission, which seem to be influenced by latitude. As already noticed in Figures 10 and 15, *An. funestus* transmission is indeed mainly concentrated in southern Katiola – possibly due to a common influencing factor within the vicinity. However, considering its low level of variance explained, it is concluded that no reasonable first or second order spatial variation model could be deduced for the malaria related data.

The drawback of regression analysis is that, although the semi-variograms tend to suggest that second order effects are insignificant in the dependent variables, most real spatial data violate this assumption – the random errors ϵ are indeed dependent and residuals are spatially correlated. It is also possible that its variance σ^2 will not be constant over the study region – a process referred to as heteroscedasticity. Therefore, although the simple least squares regression model may often provide useful insight into overall trends and relationships, results need to be treated with caution. With no clear global or local spatial trend identified, the focus

therefore shifts to the possible influence which factors within the local environment could have on the disease's erratic behaviour, as is consequently discussed.

4.3 ASSOCIATIVE ANALYSIS

Within small-area health studies any unusual local concentrations of rates or cases in time or space (as identified e.g. in probability maps) are more likely to be due to some common external factor, e.g. environmental exposures (Stern 1995:10; Mayer 1986:77). This possible association was subsequently investigated as the study's third objective – both visually and statistically.

4.3.1 Visual Identification of Associations

Differences, clustering, heterogeneity, or homogeneity within the data can be easily visualized through categorizing and the use of differently sized symbols and colours – often disclosing general associations between disease and possible determinants (Brétas 1995:90; Howe 1986:35-36; Yoon 1995:209). Proportional circle maps of fever episode and malaria transmission data, overlaid onto the villages' surrounding environmental factor observations, were consequently utilised in order to visually identify possible covariance between the dependent and independent variables (Figures 19 through 28). It should be noted that since the observations of fever episode and transmission data, within the same survey, could extend over more than one month, the independent environmental variables, as mapped within the figures, could therefore also have been observed at different months in relation to each other, but in the same month as the dependent variable under investigation.

Possible relationships among dependent and independent variables, identified through visual comparisons, are subsequently described:

- In Figures 19 and 20 it is clear that the general increase in humidity runs parallel with increases in both fever episodes and transmission intensity.
- In Figures 21 and 22 the size of villages' surrounding water bodies seem fairly consistent. No strong association could be identified.

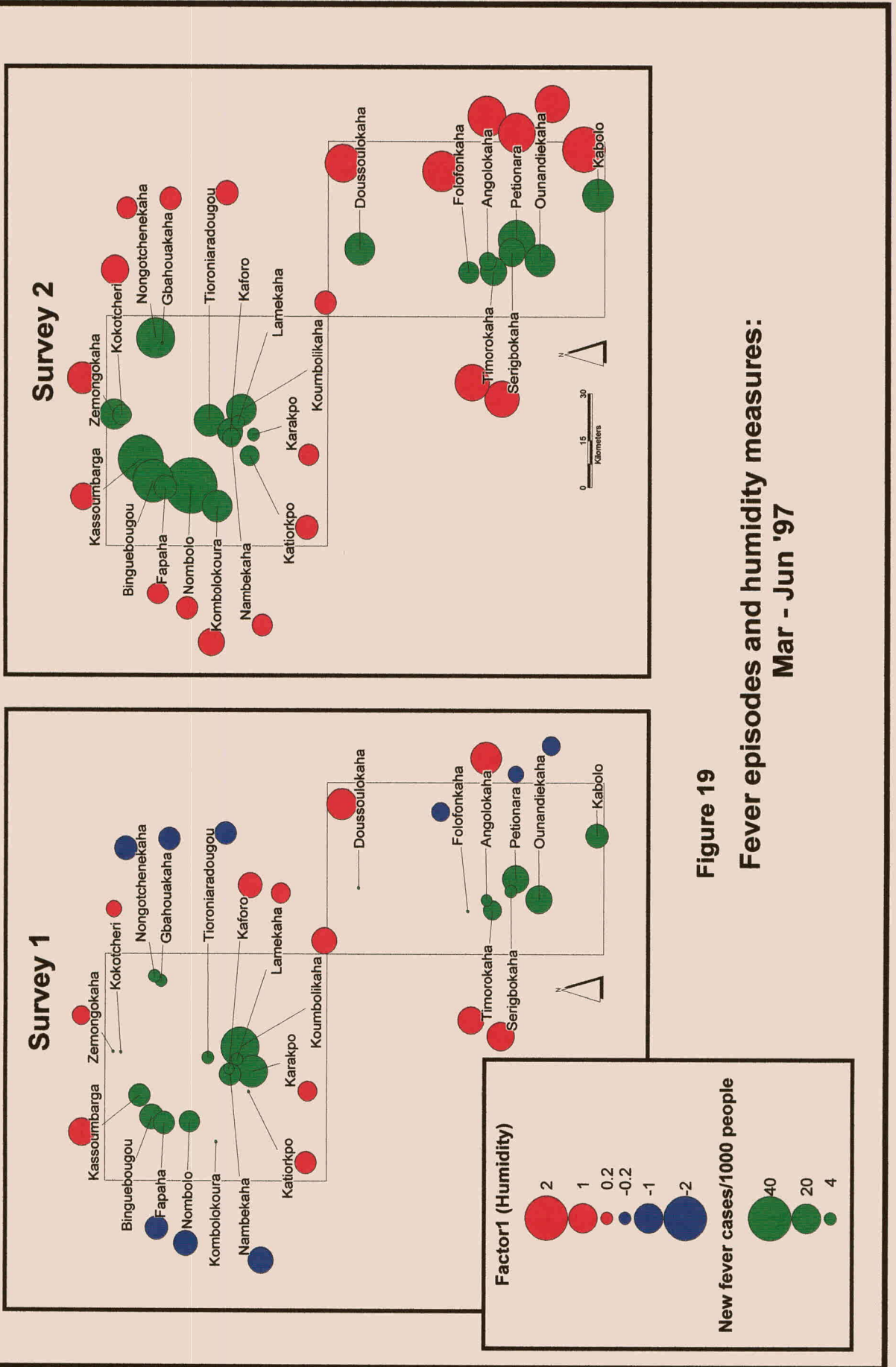


Figure 19
Fever episodes and humidity measures:
Mar - Jun '97

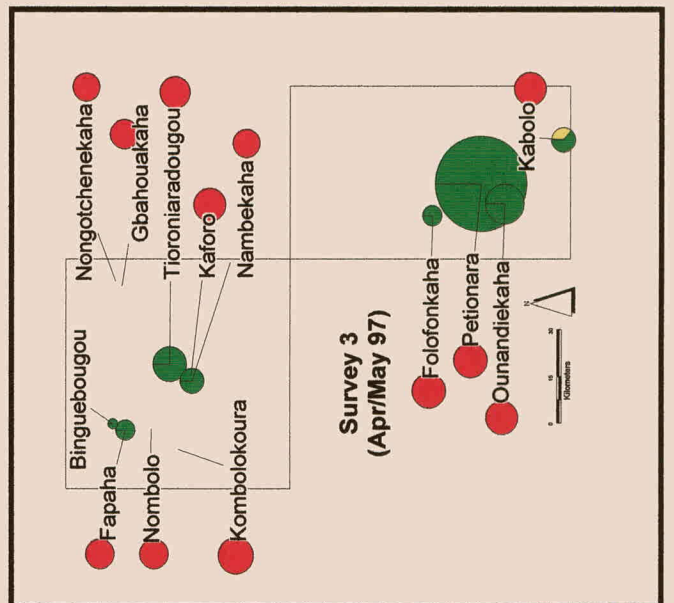
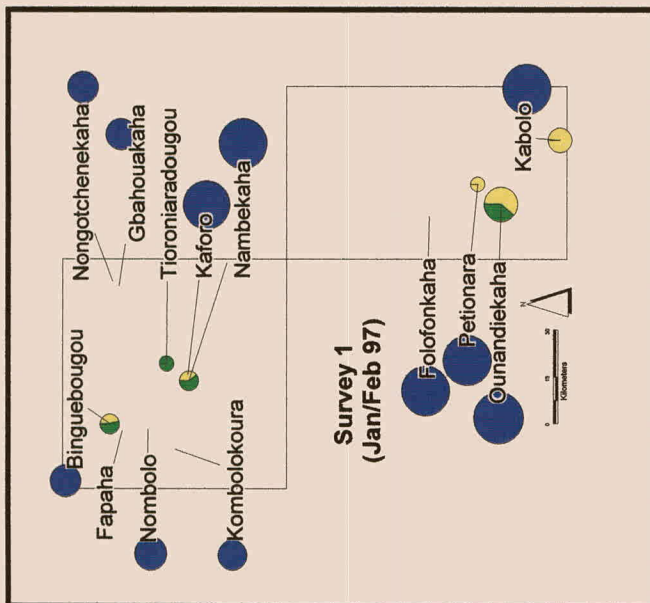
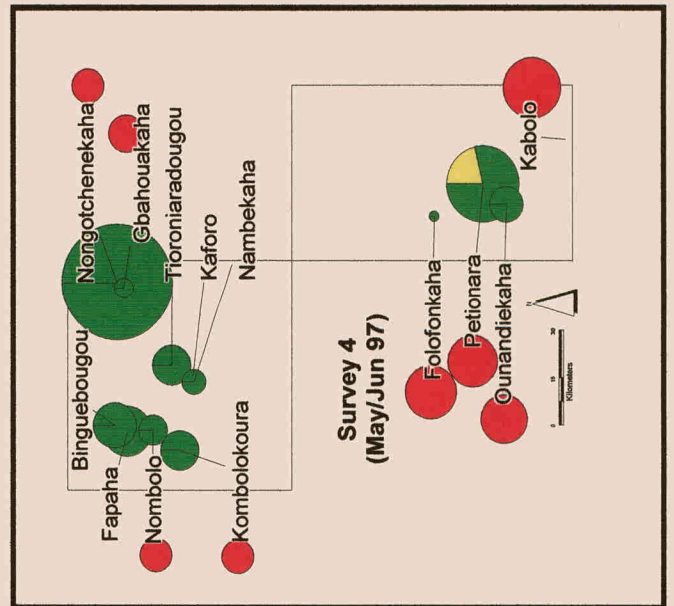
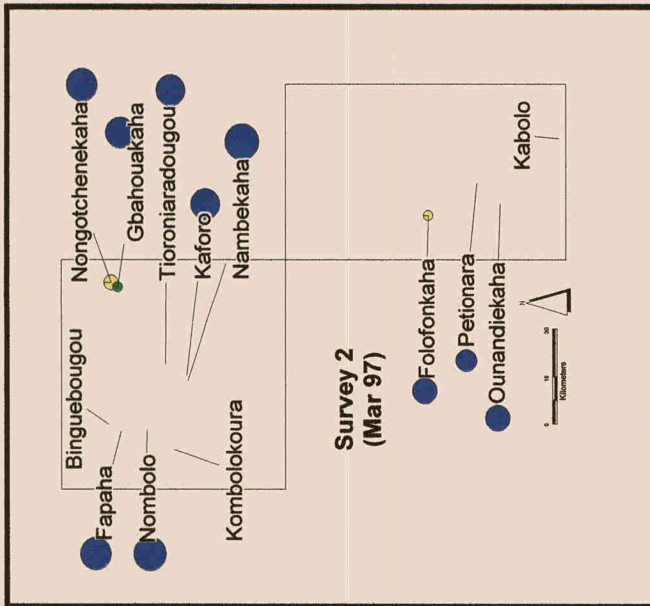
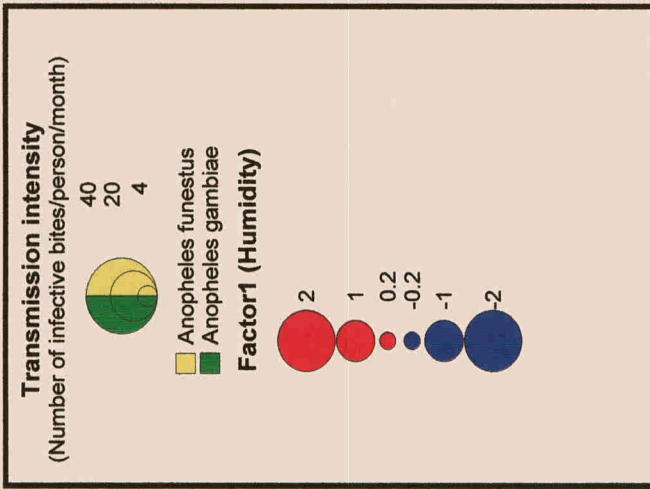


Figure 20
Malaria transmission level and humidity measures
Jan - Jun '97

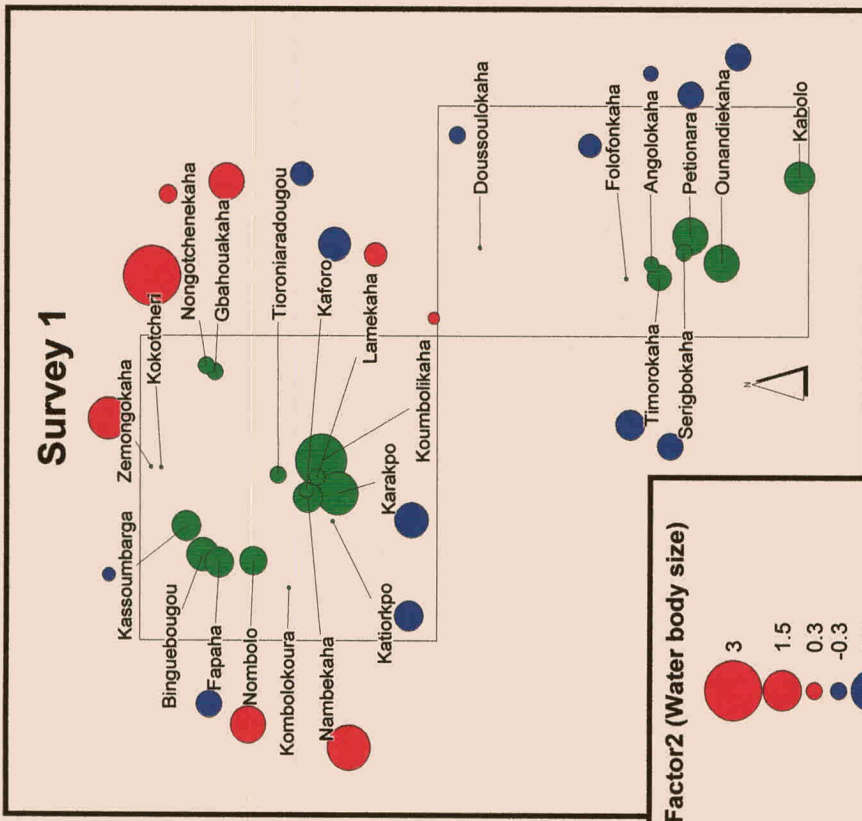
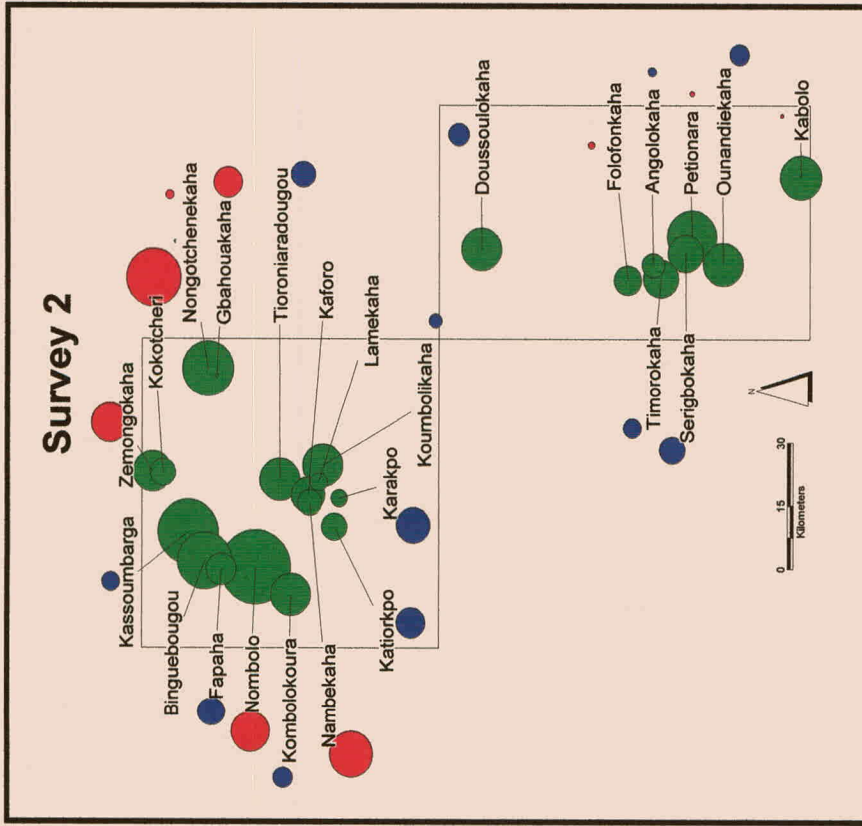


Figure 21
Fever episodes and water body size measures:
Mar - Jun '97

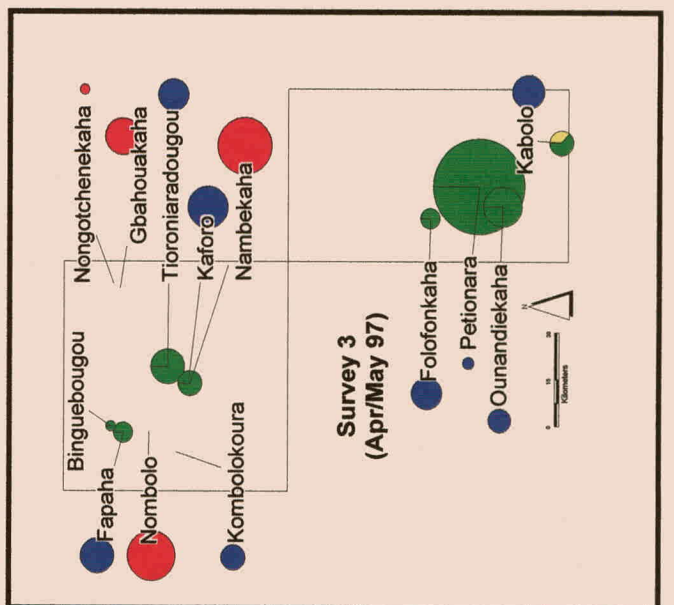
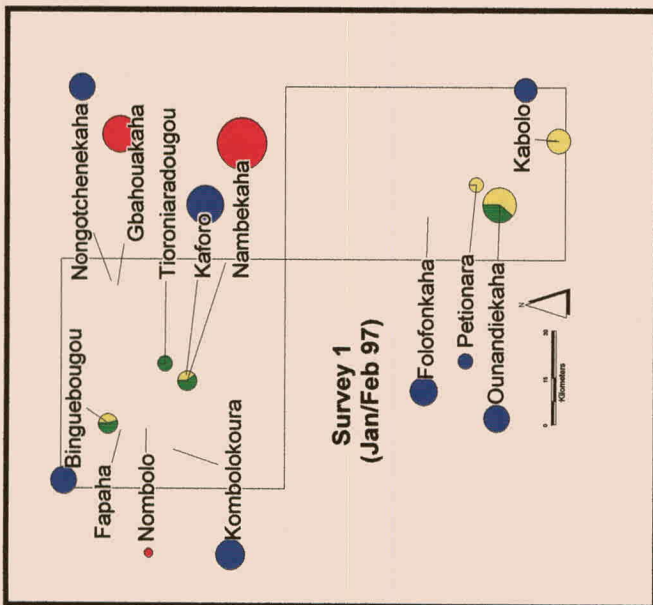
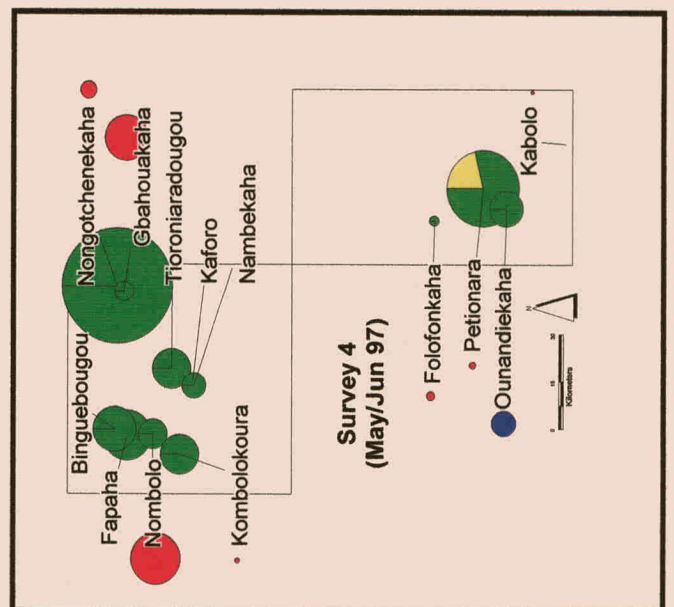
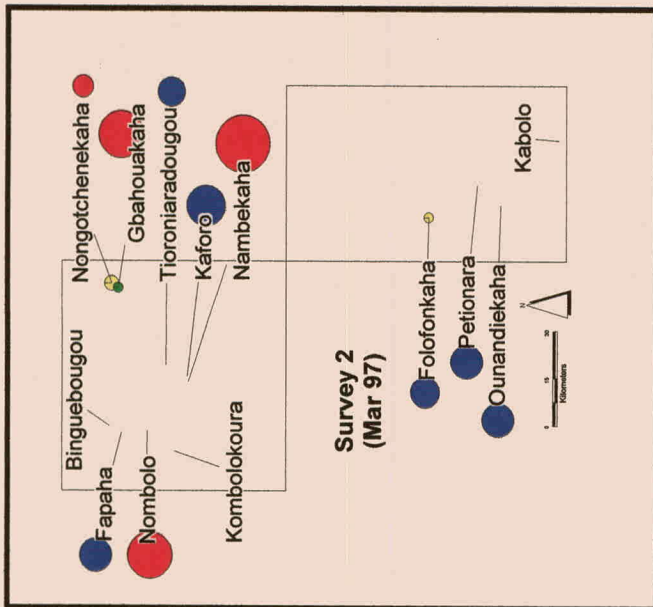
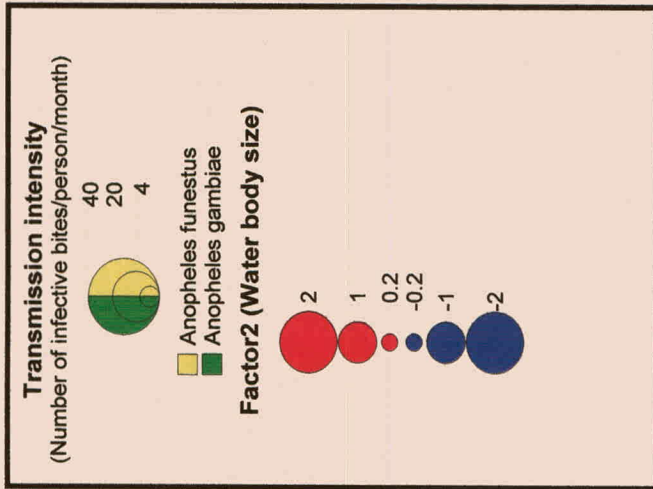


Figure 22

Malaria transmission level and water body size measures Jan - Jun '97

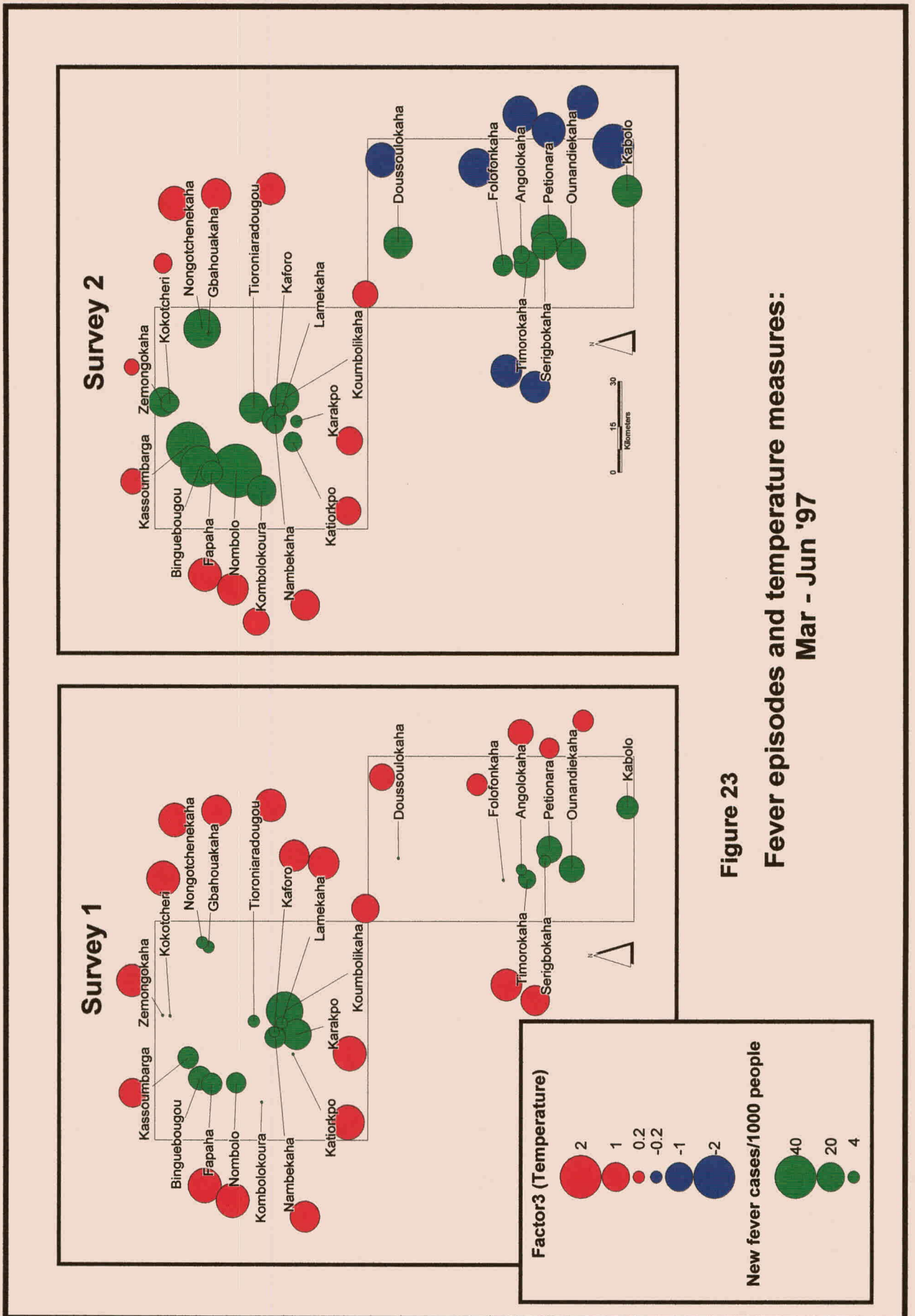
- It can be seen in Figure 23 that a decrease in temperature, especially in the Katiola district, is occurring concurrently with an increase in fever episodes. In Figure 24 it is also clear that a decrease in temperature is possibly associated with an increase in transmission intensity, and *vice versa*.
- Distance between village and water body remained relatively constant between the surveys (Figures 25 and 26), and no strong relationship could be identified.
- No specific association could be observed between the dependent variables and the level of barrenness around the villages. It is clear though that less barrenness could be observed in the Katiola district and northern parts of the study area (Figures 27 and 28).

The apparent association between a disease pattern and possible environmental determinants (e.g. humidity and temperature) might be easy to spot visually by coverage overlay in GIS, but a more robust evaluation of the pattern, as applied in the next section, is surely essential for an epidemiology GIS application (Raybould, Nicol, Cross & Coombes 1995:281-282).

4.3.2 Statistical Analysis of Associations

Using the analysis capabilities of statistical software, the associations between the spatial variations of malaria disease rates, transmission intensity and surrounding environmental influencing factors could be quantified. A simple approach to model the malaria related variables' dependence on the environmental factors is through multiple regression analysis – similar to what has been applied within Section 4.2.3. It involves the fitting of a function of the independent covariates measured at the point locations to their observed dependent variable Y by ordinary least squares regression, deriving estimates β_i and their associated standard errors. Again, standard regression assumptions, as described within the previous section, had to be assumed.

The first application concerned the observation made in Section 4.1.1, namely possible positive covariations between fever episodes and transmission levels. Forward stepwise regression was applied to these malaria risk indicators, retaining only those terms with t values that are significant at the specified α level.



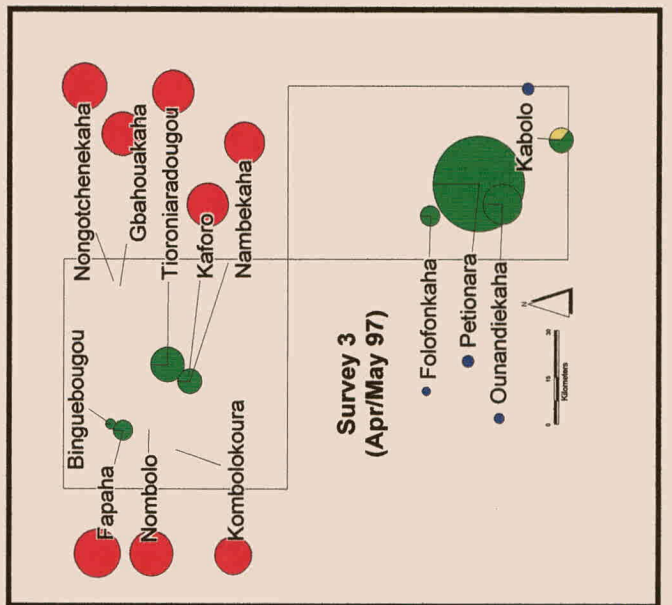
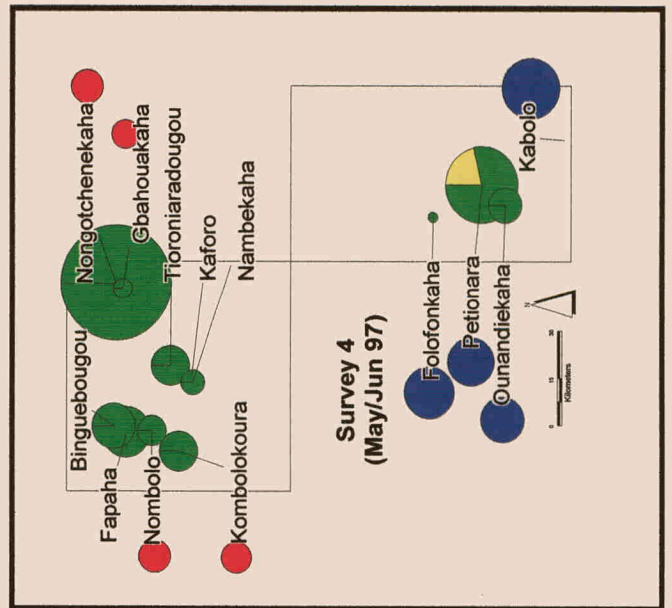
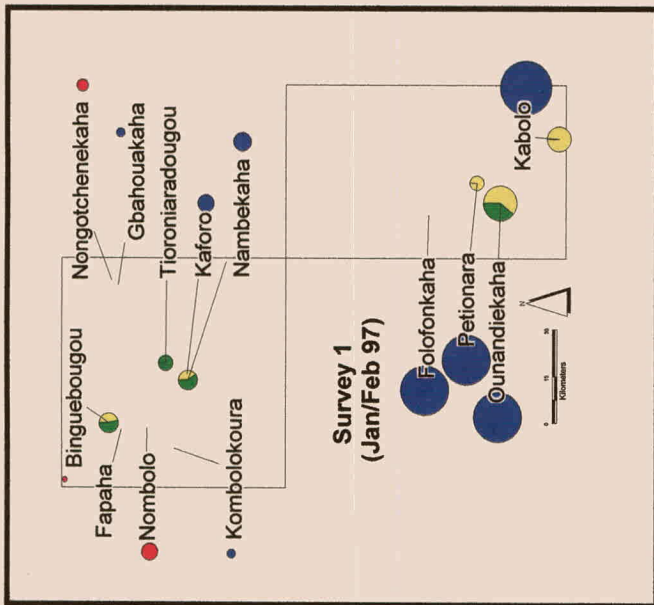
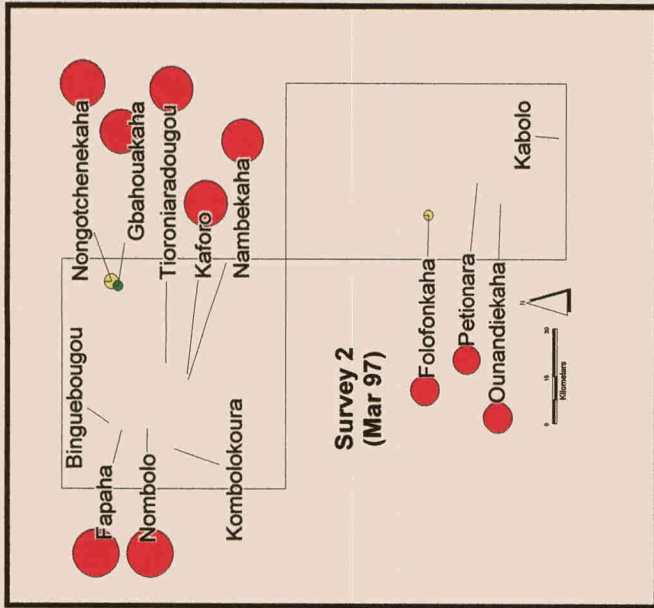
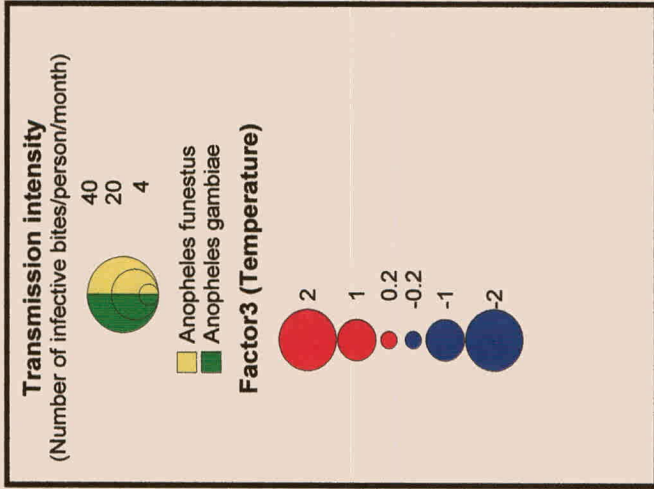


Figure 24
Malaria transmission level and temperature measures
Jan - Jun '97

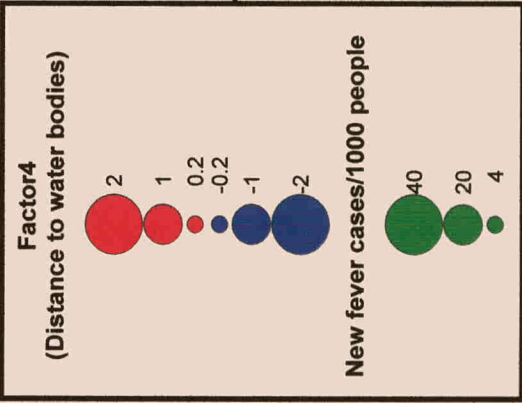
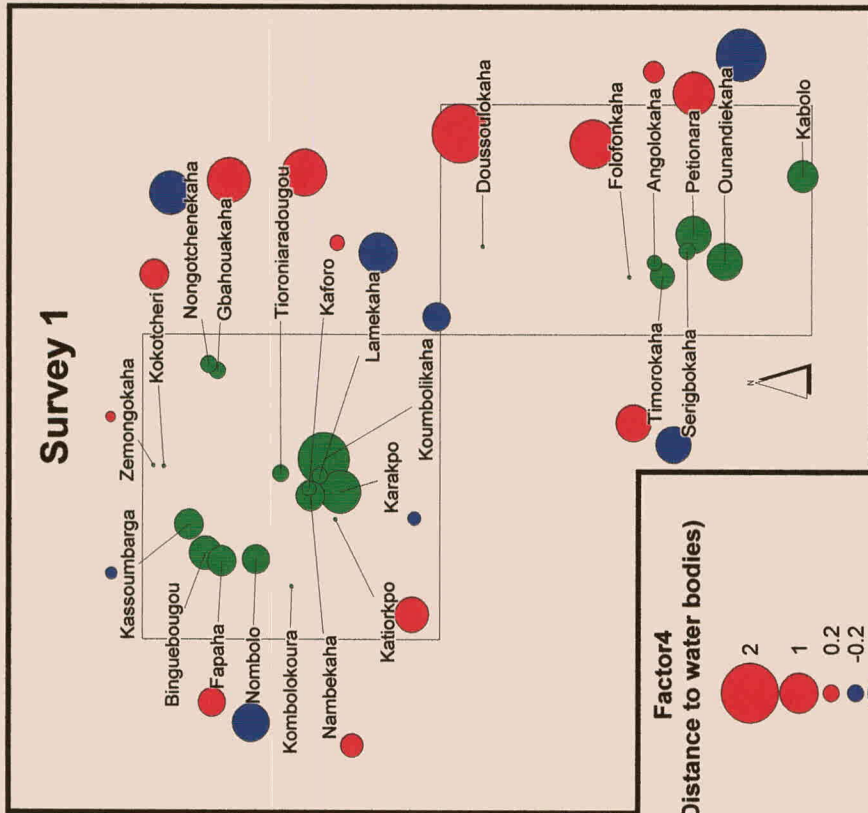
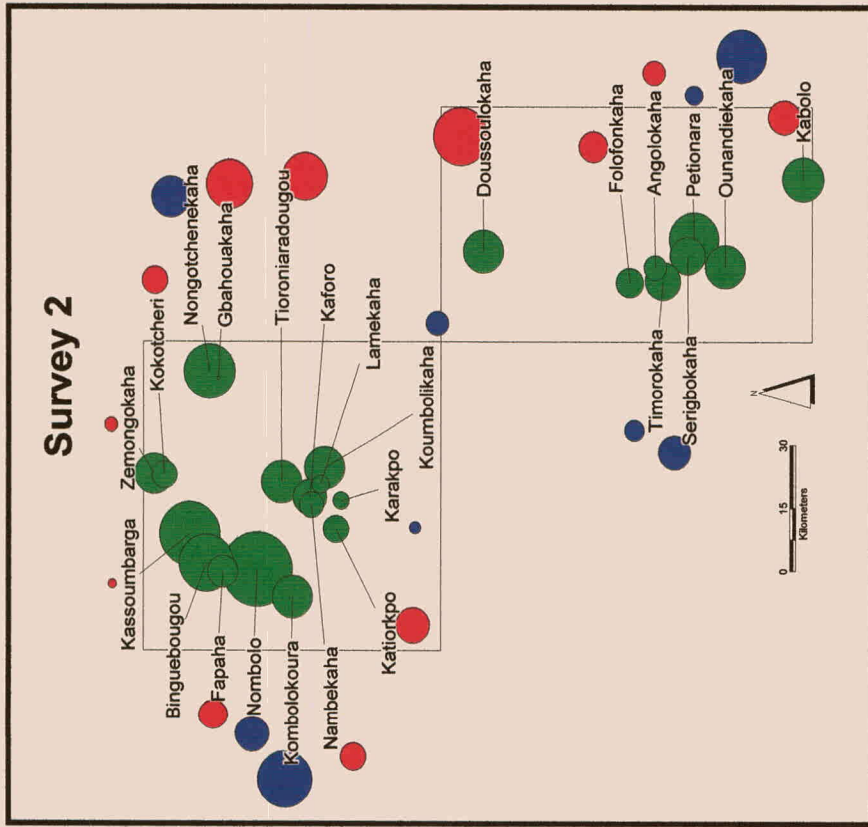


Figure 25
Fever episodes and distance measures between village and water bodies: Mar - Jun '97

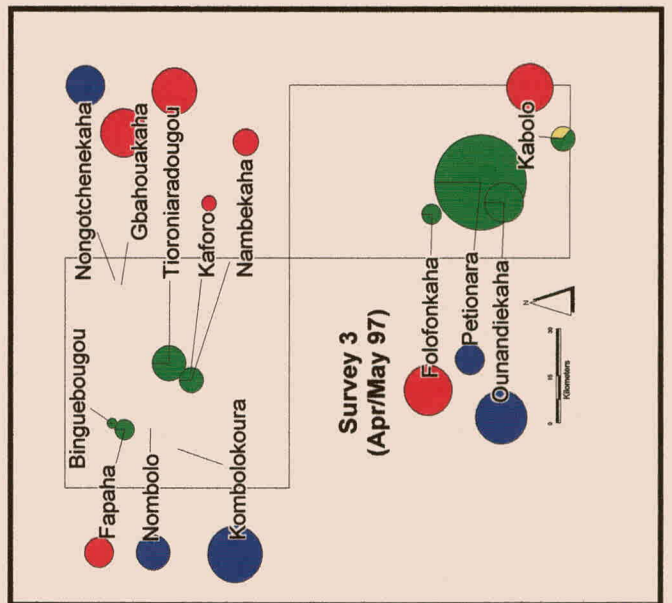
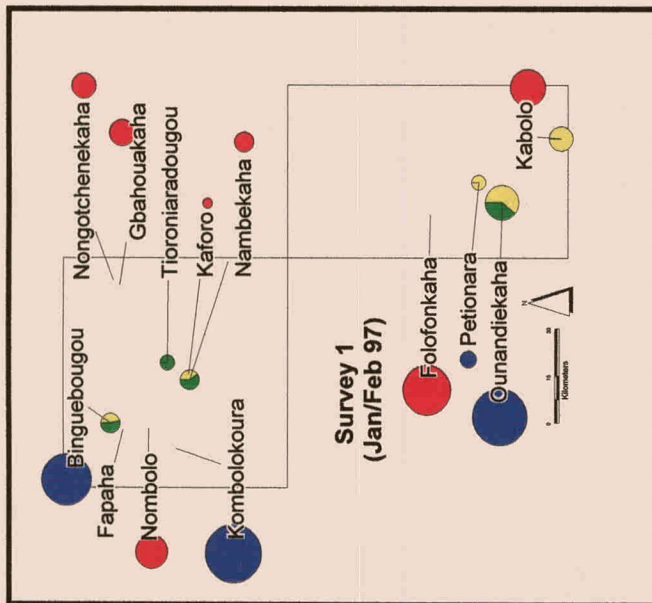
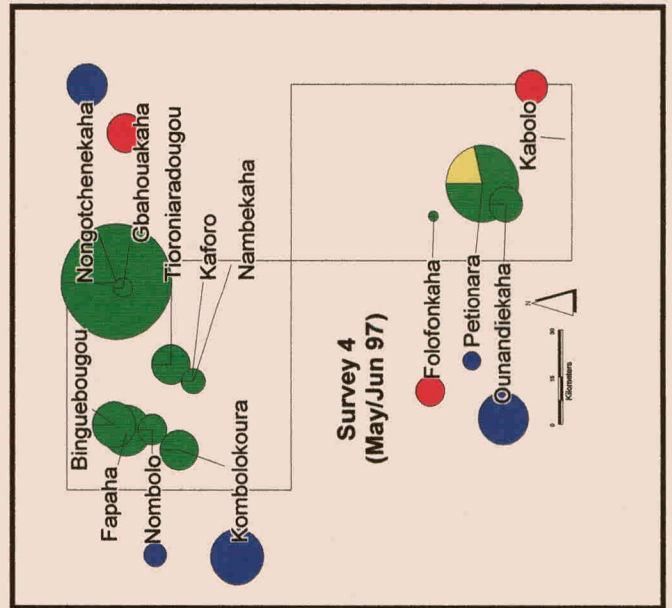
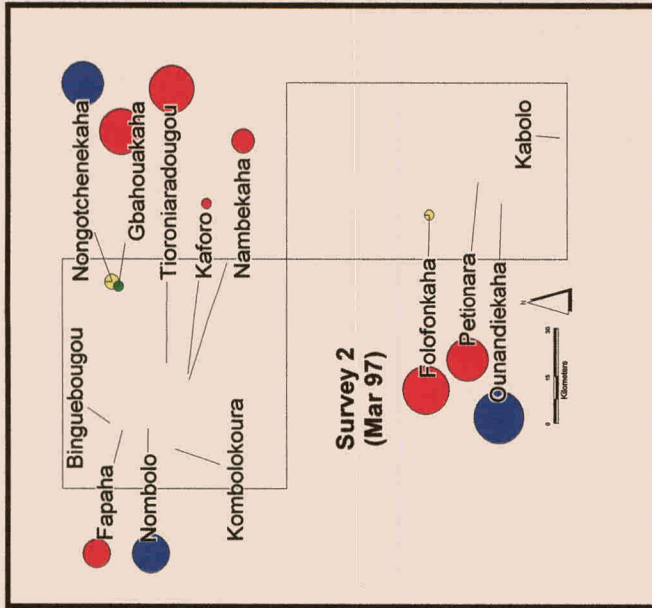
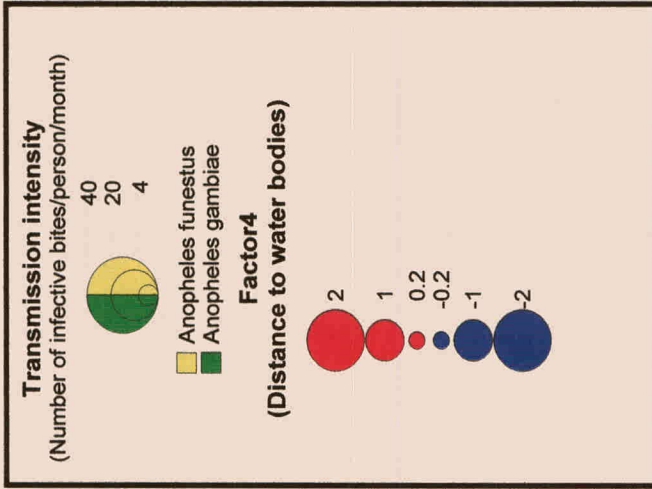


Figure 26
Malaria transmission level and distance measures between village and water bodies
Jan - Jun 97

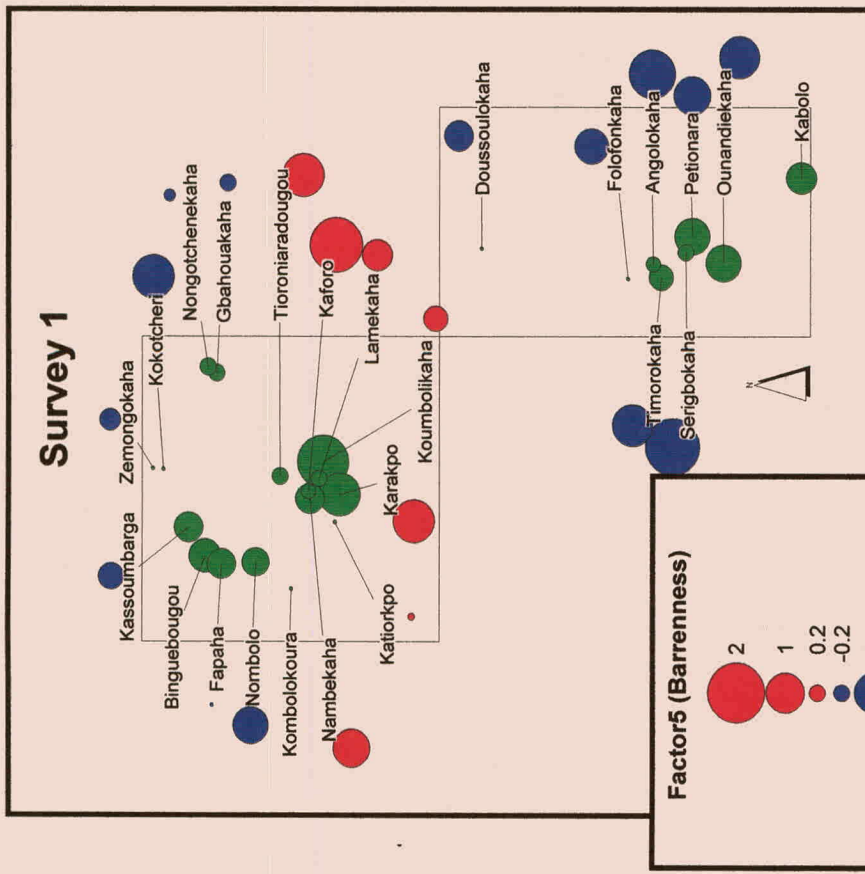
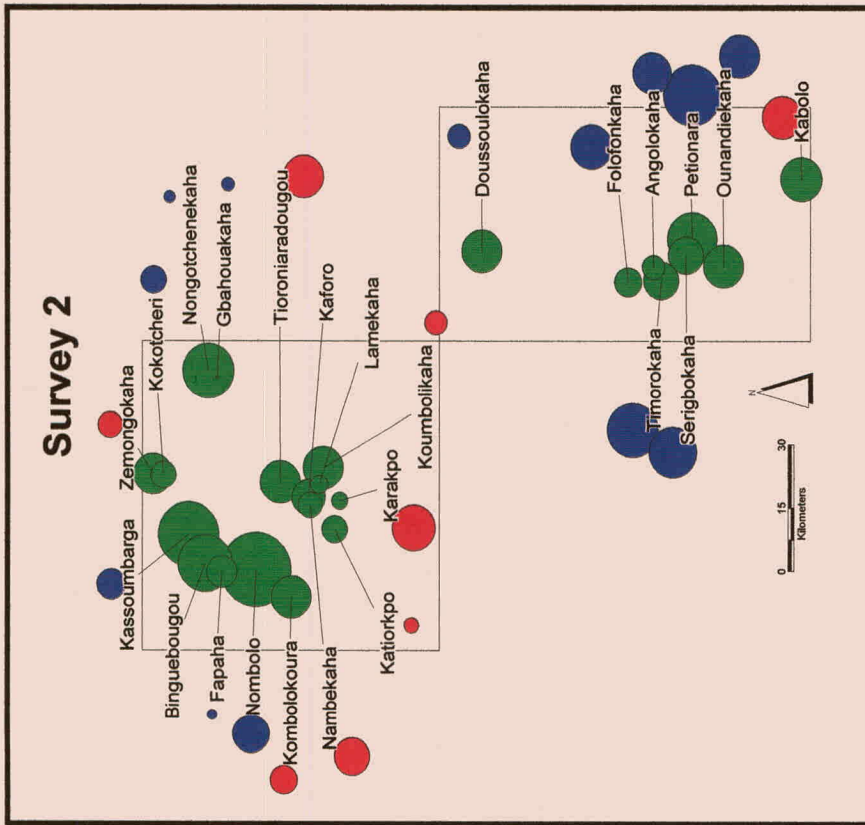


Figure 27
Fever episodes and barrenness measures:
Mar - Jun '97

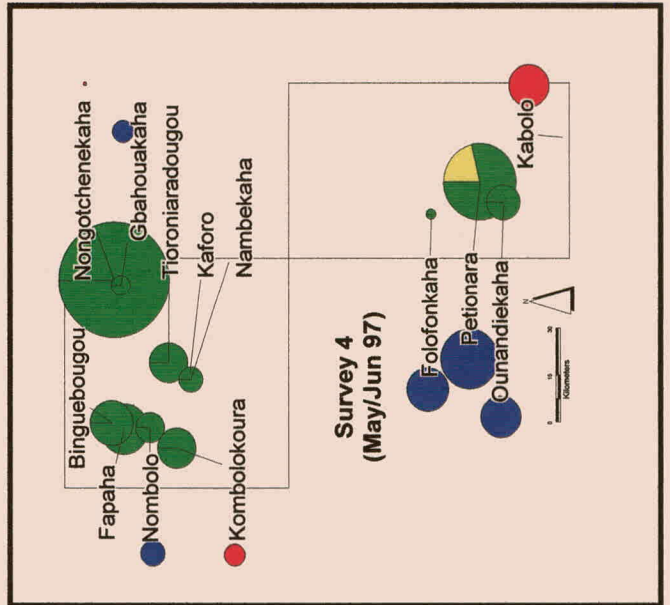
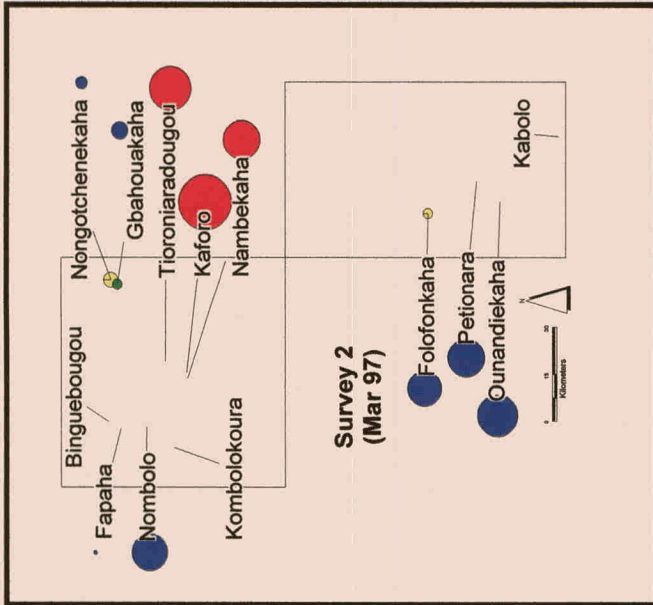
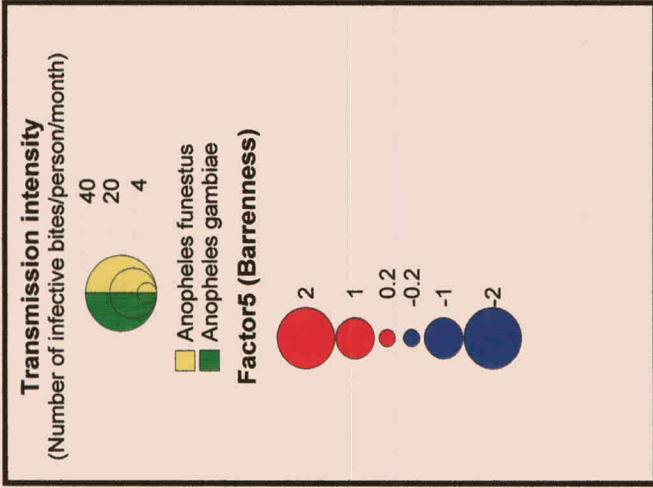
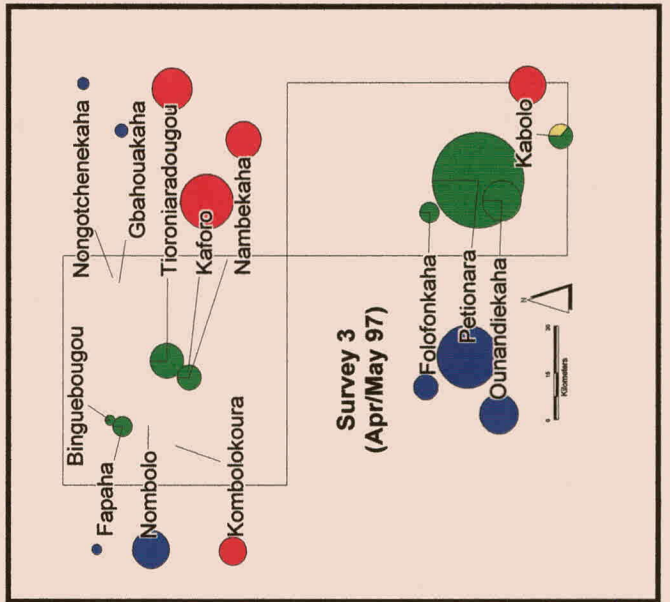
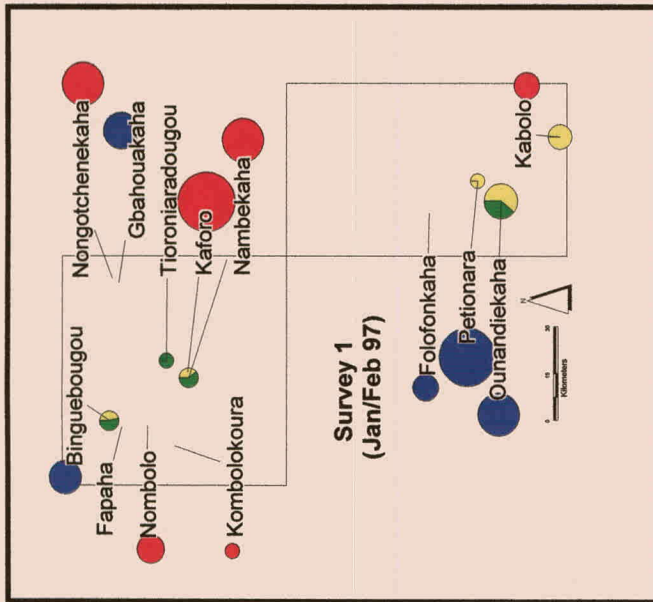


Figure 28
Malaria transmission level and barrenness measures Jan - Jun '97



The resulting regression model produced was:

$$Episodes = 14.15255 + 0.473973Ag_trans4$$

The accompanying statistics were:

	N	R ² (%)	Adj. R ² (%)	F-ratio	p	df	σ	Durbin-Watson
<i>Episodes</i>	26	6.70	2.81	1.7240	.20160	1,24	12.756	1.8241

It is insinuated that 6.7% of fever episode variance is explained by the model. However, it is clear that the model cannot be accepted since the significance level for F, namely p, is greater than an α of 0.05. Since transmission levels vary between 0.1 and 1.0 infective bite per person per year, this lack of a proportional relationship between malaria morbidity and transmission intensity is not surprising, given previous research results as sited in Section 2.3.1.

It was consequently found that a regression application of transmission and environmental data as independent variables to *Episodes* incidence would not be feasible, since the number of valid cases only amounts to 21. Further analysis therefore entailed environmental factors to be the only independent variables, and fever episodes and transmission risk to be the dependants. Forward stepwise regression was applied to these combinations, with resultant regression models as follows:

$$Episodes = 12.29288 - 4.781689Factor4 + 3.834439Factor1$$

$$Af_trans = 0.977166 - 0.833577Factor3 - 0.340901Factor1 - 0.327095Factor5$$

$$Ag_trans = 6.728022 + 5.52968Factor1 - 3.294961Factor5 - 2.763895Factor4$$

The accompanying statistics were:

	N	R ² (%)	Adj. R ² (%)	F-ratio	p	df	σ	Durbin-Watson
<i>Episodes</i>	42	18.0	13.8	4.2829	.02082	2,39	11.425	1.4727
<i>Af trans</i>	41	28.9	23.2	5.0219	.00508	3,37	1.7009	1.7720
<i>Ag trans</i>	41	18.8	12.2	2.8569	.05010	3,37	16.821	2.5515

It is implied that the amount of variance explained by the various models are respectively 18% for fever episodes, about 29% for *An. funestus* transmission, and almost 19% for *An. gambiae* transmission. Using the F-test, the null hypothesis of no explanation ($\beta_1 = \beta_2 = 0$) is, however, accepted for the *Ag_trans* regression model at an α of 0.05. Evaluating the residuals for autocorrelation using the Durbin-Watson statistics, it is concluded that both the *Episodes* and *Ag_trans* models cannot be accepted due to significant autocorrelations.

The interpretation is that most of the observed y values lie within 2σ of the least squares predicted values, therefore within: ± 22.85 , ± 3.40 , and ± 33.64 respectively. Since the R^2 's of all models are very low, the residuals were plotted against the model estimates, and the outliers with a residual $> \pm 2\sigma$ identified. Comparing these with earlier Figures 9 and 10, it is clear that the outliers of *Episodes* (i.e. Nombolo and Kassoumbarga within Survey 2) form a “hot spot” in the north west of the study region. The outlier of *Af_trans* (i.e. Petionara, Survey 4), however, does not seem unusual compared to the transmission intensity of *Ag_trans*. The outliers of *Ag_trans* (i.e. Petionara, Survey 3, and Nongotchenekaha, Survey 4), on the other hand, seem very large compared to the norm. By visually comparing the outliers with their accompanying environmental factors, no apparent reason can be put forward for their existence or location – probably due to the influence of other external factors not included in the study. Data was verified, and chances of observational errors are small. The outliers of *Episodes* and *Ag_trans* were therefore temporarily omitted from the regression procedure, which led to significantly improved R-squares, F-ratios, and more stable models than those with outliers included.

The models generated were:

$$Episodes = 15.4691 - 4.088695Factor3 - 3.845461Factor4 + 2.406491Factor5$$

$$Af_trans = 0.977166 - 0.833577Factor3 - 0.340901Factor1 - 0.327095Factor5$$

$$Ag_trans = 3.658997 + 3.039363Factor1 - 1.173378Factor5 - 1.053615Factor3 - 0.854143Factor4$$

The statistics were:

	N	R ² (%)	Adj. R ² (%)	F-ratio	p	df	σ	Durbin-Watson
<i>Episodes</i>	40	31.1	25.4	5.4277	.00348	3,36	8.0142	1.4742
<i>Af_trans</i>	41	28.9	23.2	5.0219	.00508	3,37	1.7009	1.7720
<i>Ag_trans</i>	41	33.7	25.9	4.3181	.00624	4,34	5.5989	1.8108

The amount of variance explained by the *Episodes* model has therefore been increased from 18% to 31%, and that of *Ag_trans* from 18.8% to almost 34% by excluding their respective outliers. Although the null hypothesis of no explanation ($\beta_1 = \beta_2 = 0$) is rejected for all three models at an α of 0.05, the Durbin-Watson statistic indicated that the *Episodes* model could still not be accepted due to a significant autocorrelation between residuals.

Investigating the possibilities of polynomial functions, it was found that no second or third order polynomial regression could significantly improve the *Episodes* model. The *Af_trans* and *Ag_trans* regressions could, however, be improved by including second and third order terms, as are consequently presented. The largest outlier with a standard residual $> \pm 2\sigma$ was again omitted from the *Ag_trans* model:

$$Af_trans = -0.0916 + 0.970567(Factor1)^2 - 0.233927(Factor5)^3 - 0.278488(Factor1)^3$$

$$Ag_trans = -1.42849 - 3.523286(Factor5)^3 + 3.884662(Factor5)^2 + 4.241388Factor5 + 1.529854(Factor4)^2$$

The accompanying statistical findings were:

	N	R ² (%)	Adj. R ² (%)	F-ratio	p	df	σ	Durbin-Watson
<i>Af_trans</i>	41	36.1	30.9	6.9620	.00079	3,37	1.6131	1.8193
<i>Ag_trans</i>	40	61.1	56.6	13.717	.00000	4,35	7.8769	1.7901

These models, with about 36% and 61% of *Af_trans* and *Ag_trans* variance explained respectively, could be accepted above previous results since the null hypothesis of no explanation could be rejected (F-ratios were larger than critical F's), and autocorrelations between residuals were insignificant.

It was further proposed to investigate the possibility of time lags between environmental cause and disease related response. Forward stepwise regression was applied to the disease data, in relation to environmental measurements of one to three months earlier. The only model improving previous results were that of *Episodes*, with a two month lagged response:

$$Episodes = 13.33182 - 4.282849Factor4 + 3.490095Factor3 - 2.341345Factor5$$

The associated statistics were:

	N	R ² (%)	Adj. R ² (%)	F-ratio	p	df	σ	Durbin-Watson
<i>Episodes</i>	42	23.0	17.0	3.7923	.01791	3,38	11.663	1.8549

This model is consequently accepted above previous results, since the null hypothesis of no explanation could be rejected, while the Durbin-Watson statistic indicated insignificant autocorrelations amongst residuals. Neither exclusion of outliers or inclusion of polynomial terms could significantly improve the 23% variance explained.

These fitted lines or polynomials in themselves, however, are not of direct relevance in the search for associative relations. More relevant are the residuals: the differences between the observed data values and the regression function. The mapping thereof can give an indication of associative relations: positive residuals occur in areas where death rates are underestimated by the regression model, and negative residuals represent districts with overestimated morbidity (Douven & Scholten 1995:128; Cliff & Hagget 1988:47). The residuals as obtained from the final lagged *Episodes* model and the final third order polynomial transmission models are portrayed in Figures 29 - 31. It can be seen that the residuals obtained from the *Episodes* model are quite high, and that no clear pattern or clustering of same-order residuals could be identified. The residuals of the *An. funestus* model are quite low in the second and third survey, mainly due to the low level of transmission that took place in those months. In contrast with the R² of about 61% obtained in the *An. gambiae* model, large residuals are observed in some instances, varying in no consistent manner over the surveys. The actual suitability of the model is therefore questioned.

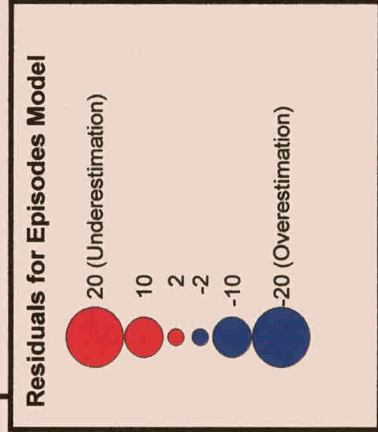
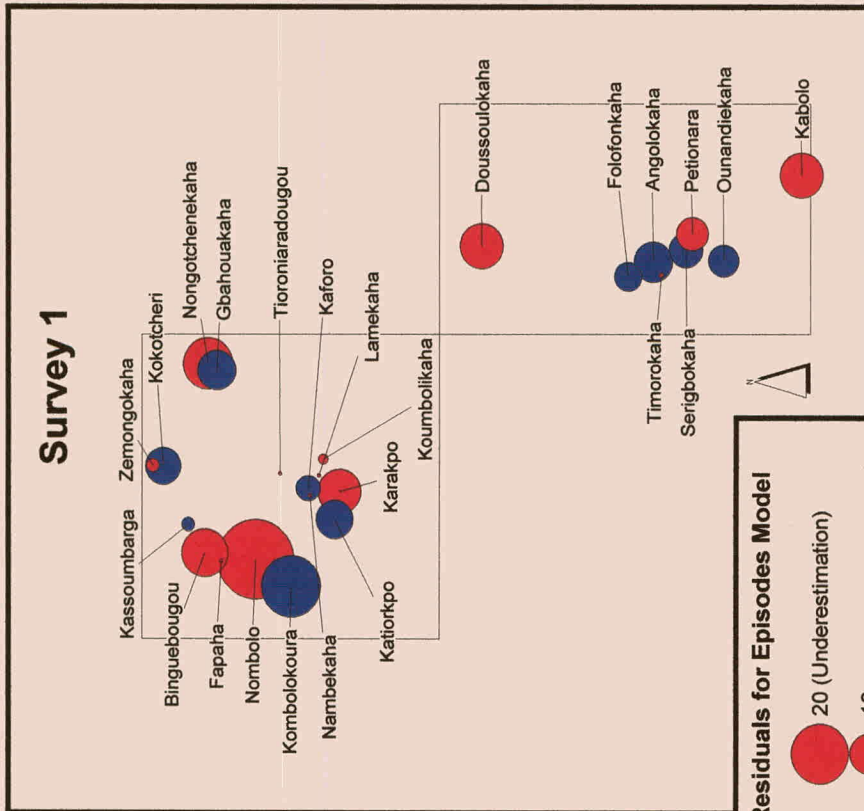
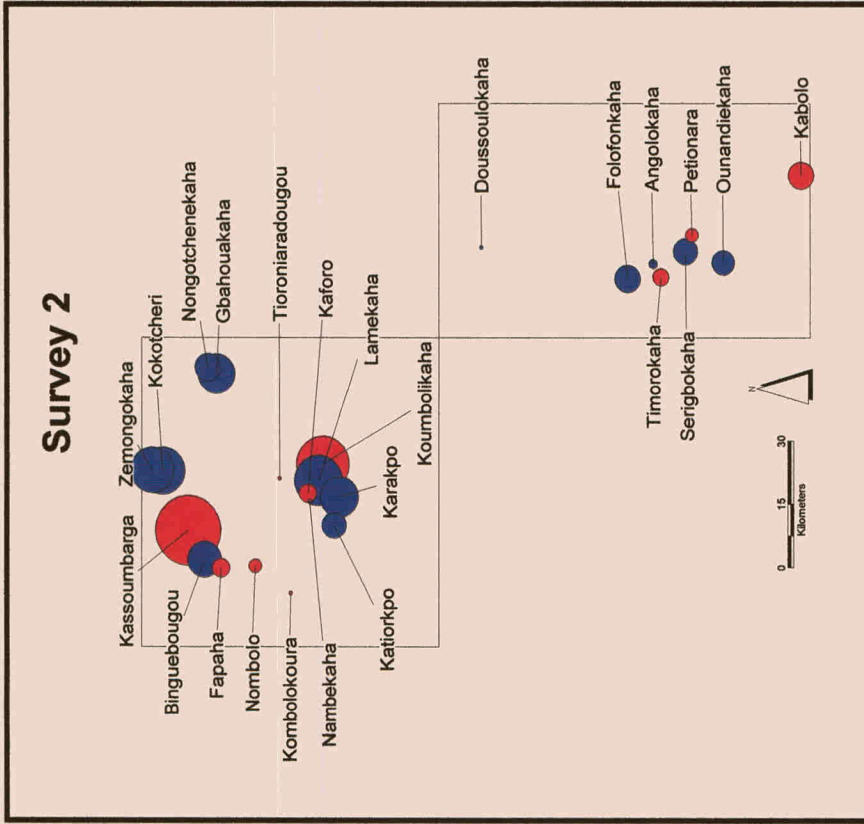


Figure 29
Residuals obtained from the fever episode model

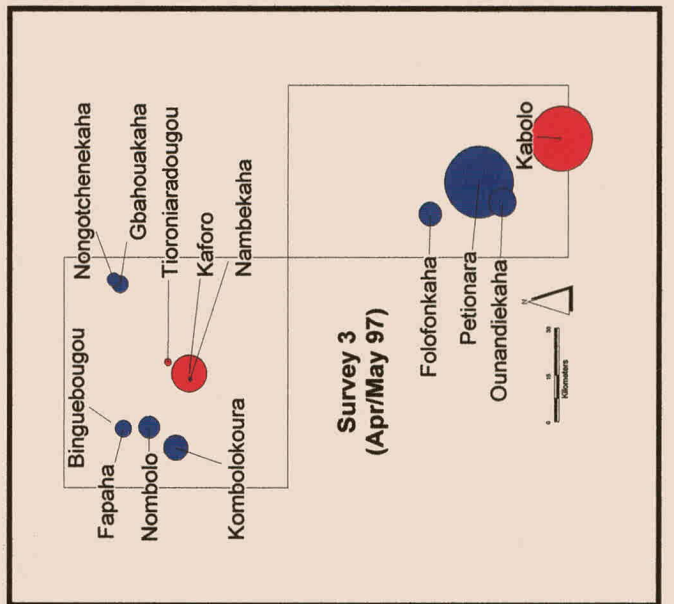
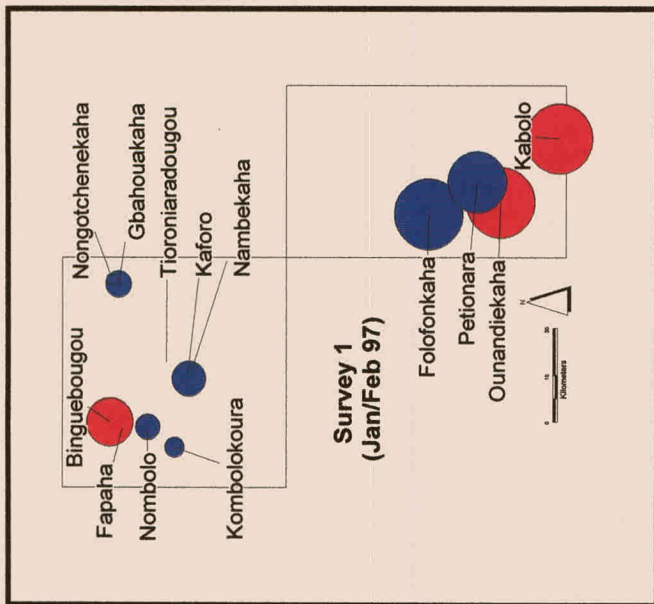
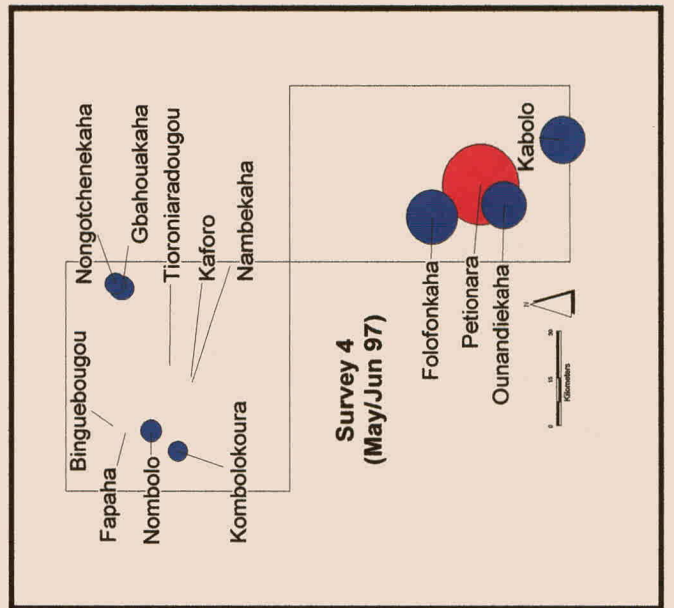
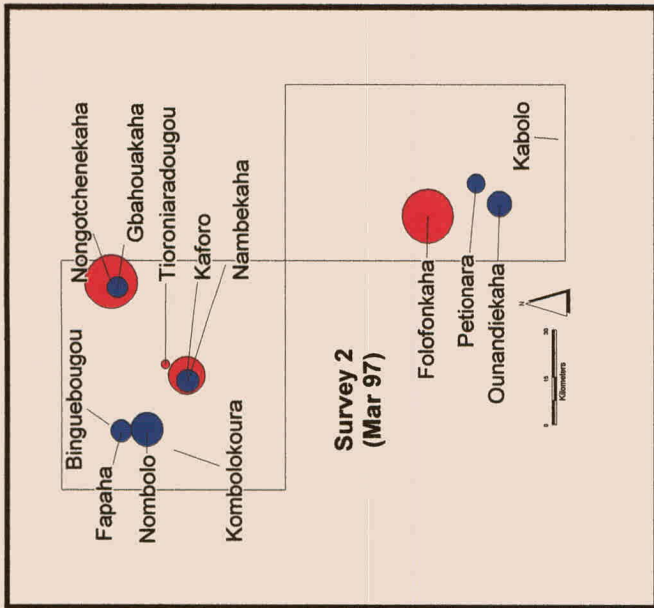
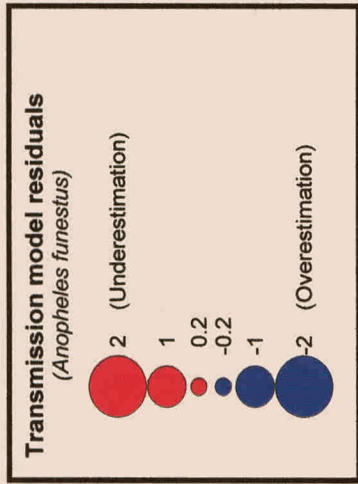


Figure 30
Residuals obtained from the
Anopheles funestus
model

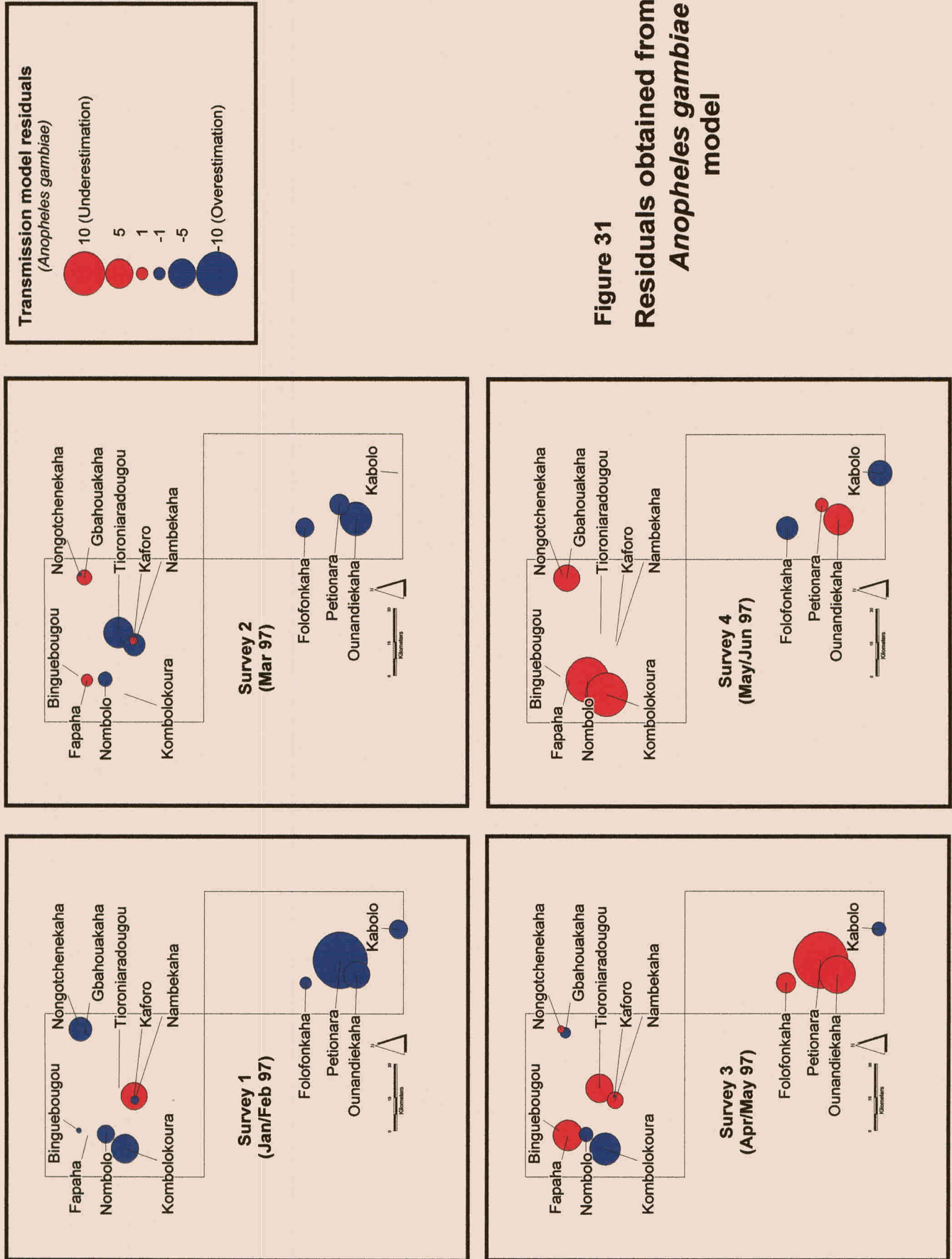


Figure 31
Residuals obtained from the *Anopheles gambiae* model

Due to the small proportion of the variance in the dependent variables accounted for by some of the regression models, and to the high variability in residuals in others, the models were not used as inputs to statistical modelling of disease incidence in regions where field data is unavailable. Although a substantial degree of morbidity and transmission could be accounted for by measures of relative humidity, temperature, distance to water bodies and vegetation density, it is clear that other determinants still need to be identified and studied in order to establish stable and reliable predictive models. It is interesting to note that *Factor2*, namely water body size and its flooding pattern, does not make a significant contribution to any of the regression models. A re-evaluation of the necessity and relevance of painstakingly measuring the amount of flooding within wetlands should therefore be done, to ensure the optimal allocation of time and finances to measurements of possible new factors. It is suggested that attention should be given to the areas or villages prone to high levels of fever episodes and transmission (i.e. in the north western part of the study area, Petionara and Nongotchenekaha), in order to identify additional factors possibly responsible for the high incidence.

4.4 CONCLUSIONS

The focus within this chapter was on visual and statistical investigations of malaria's spatial patterns, variations and associations with environmental influencing factors, employing a method known as exploratory spatial data analysis. Within this analysis, Geographic Information Systems (GIS) played a meaningful role with regards to:

- mapping of malaria distribution, extreme values and regression model residuals,
- exploration of first order variation through TIN interpolation, and
- the study of second order variability by means of semi-variograms.

Statistical packages were used in:

- identification of extreme malaria occurrences, through detection of outliers exceeding the 90th percentile of a normal probability distribution,
- modeling first order variation through trend surface analysis,
- investigation of associations with environmental influencing factors through standard, polynomial and time-lagged regression modeling, and
- the computation of regression model residuals.

Within the final chapter a short summary of research results is given, including some suggestions for further research.

CHAPTER 5: CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Malaria, being the most important vector-borne disease in Africa, is still leaving an ever-increasing trail of sickness and death within the developing world, not to mention its economic impediments on development and productivity. Resulting in 270-480 million clinical cases annually in sub-Saharan Africa alone, it is also rated as Africa's single largest killer of children – approximately one million per annum. Surprisingly, although it is known that malaria's spatial distribution and intensity are influenced by numerous environmental variables, little research has been done in modelling the disease risk for different geographic locations as based on these environmental determinants.

Medical geography's study of the spatial human-environmental interrelationships of disease can make a significant contribution to the understanding of malaria's spatial variations in transmission and morbidity, and its interaction with environmental influencing factors. This study consequently proposed to investigate the nature of the disease's spatial variation within a sample frame of 24 villages in northern Côte d'Ivoire, being representative of rice cultivation areas within West Africa. The investigation's objectives included the development of malaria risk maps, an analysis of its spatial variation through exploratory spatial data analysis and GIS functionality, the search for and application of predictive models as based upon environmental variables, and an assessment of the accessibility of health services responding to this anticipated health risk.

In order for relevant models to be derived, the malaria cycle, which involves the transmittal of the *Plasmodium* parasite between anopheline mosquito and humans, had to be reviewed. The dominant mosquito carriers (vectors) of the parasite within the study region consist of the *An. funestus* and *An. gambiae* species – each with a unique array of breeding preferences and responses to environmental influences. Malaria transmission intensity has been studied as a factor having a direct influence on the malaria transmission cycle, and indirect factors included in the study were:

- wetland rice production which provide extended and prolonged mosquito breeding habitats;

- water bodies and their flooding pattern, providing the aquatic environment for vector breeding;
- temperature ranges which influence vector and parasite mortality;
- relative humidity which enhance vector longevity if sufficiently high;
- amount of rainfall, regulating the availability of breeding habitats;
- insolation, influencing the permanence of breeding pools and the ranges of temperature and relative humidity;
- distances between aquatic breeding grounds and villages – the source of human blood; and
- vegetation cover, serving as resting places, flight barriers and home to additional breeding puddles, as often expressed in the Normalized Difference Vegetation Index (NDVI).

Malaria related data subsets have been obtained from surveys carried out by Institut Pierre Richet/OCCGE within the 24 study villages during the dry season of 1997. Environmental information have been acquired from measurements by the WARDA Health Research Consortium around these villages, district level meteorological observations, and GIS computations of distances and vegetation biomass measures using satellite imagery. Data obtained for the analysis consisted of three malaria risk indicators, namely fever episode occurrences, *An. funestus* and *An. gambiae* transmission measurements, as well as 17 measures of environmental factors which might have an influence on the disease's distribution and intensity. The original 17 independent factors were manipulated, standardised and reduced to five uncorrelated principal components through principal component analysis, namely humidity, water body size, temperature, distance to water bodies and vegetation biomass measures.

The incorporation of environmental factors into a predictive model of malaria morbidity is often restricted due to large amounts of data to be processed. By its ability to integrate large digital datasets on the basis of their spatial characteristics, Geographic Information Systems (GIS) has been found to enhance the feasibility of studying the environmental determinants of malaria. Able to rapidly retrieve, update, query, analyze, model and map the diverse data sources, it can provide timely and pertinent information, leading to more successful intervention strategies and the effective allocation of limited resources.

In order to address the study's objectives efficiently, a process known as exploratory spatial data analysis has been utilized to investigate the spatial characteristics of malaria occurrence, and its association with the above-mentioned environmental related factors. This entailed the undertaking of a series of exploratory steps, as identified by Bailey & Gatrell (1995:21-25) and Douven & Scholten (1995:119), of which a short summary of results is subsequently given:

- a) In order to identify spatial disease patterns, the malaria related factors' geographical distribution was mapped by means of proportional circle maps. It was observed that both morbidity and transmission increased during the study period, climaxing at the beginning of the rainy season. Areas of extreme disease occurrence were also identified in the north west of the study area, while Petionara and Nongotchenekaha seem to be centres of exceptional transmission risk.
- b) Intending to explore the global (first order) spatial variation in the disease morbidity and transmission variables, Triangulated Irregular Networks (TINs) were produced and values interpolated at all locations in the study area. The "hot spots" of extreme disease measures were again apparent. The local (second order) variations were subsequently examined by means of semi-variograms, which revealed a high degree of spatial variability, with little or no spatial dependence between observations. Since no clear second order variation could be identified, little would be attained to model local spatial variation through techniques such as Ordinary Kriging. The focus therefore shifted to modeling global (first order) variation through a polynomial expansion of the villages' geographical coordinates by least squares regression – a technique known as trend surface analysis. However, only a low level of disease transmission variance could be accounted for by the villages' spatial locations, and it was concluded that neither a first or second order spatial variation model could be successfully deduced.
- c) Since the disease's erratic spatial behaviour in morbidity and transmission levels was most likely to be due to local environmental influences, these possible associations were examined – first visually, and then statistically. Through visual comparisons on

proportional circle maps, humidity and temperature measures have been identified to covary with disease occurrence, therefore being possible determinants. Associations were subsequently tested statistically through standard multiple regression modeling. However, since the models could account only for low percentages in the dependent variables' variance, outliers were identified and excluded, resulting in improved models. The inclusion of second and third order polynomials, and the investigation of time lags further improved results, so that the final percentage variance accounted for by the environmental factors were 17% of fever episodes, 36% of *An. funestus* transmission and 61% of *An. gambiae* transmission. Environmental factors included in the study were relative humidity, temperature, distance to water bodies and vegetation density. The residuals of these final models have been mapped to investigate possible patterns in over- and underestimation, but no clear patterns or clusters could be identified.

- d) Since some of the final regression models explained such a small percentage of variance, and in others large residuals could be observed between actual and estimated observations, the application thereof for further malaria risk prediction in new areas, as the study's fourth objective, seemed inappropriate. Relationships between malaria morbidity or transmission and environmental factors are clearly complex and often highly site specific, possibly requiring higher order polynomial functions and a wide spectrum of determinants. In order to establish more stable models, other relevant environmental factors still need to be identified and studied – especially in villages identified with extreme levels of malaria risk. It was observed that water body size and its flooding pattern, for example, did not make a significant contribution to any of the models, and it might be worthwhile to investigate and measure other aspects of these features, such as water depth and temperature. In addition to possible environmental variables, the important influence of socio-cultural, economic and behavioural aspects should also be considered.
- e) According to Trape & Rogier (1996), excessive mortality is to be attributed to a lack of health services and medicines. However, since malaria risk could not be predicted with a high level of certainty using the generated models, the sufficiency of current medical facilities responding to malaria related risk could not be assessed, as initially intended with

the study's fifth objective, either. It would, however, be wise to target the identified malaria risk "hot spots" with the necessary medical resources.

Data subsets available for this research included surveys for only four months in the case of fever episodes, and six months with regards to malaria transmission. Additionally, some environmental data sets (e.g. rainfall measures) had to be enhanced through interpolation techniques, while some meteorological data sets (namely relative humidity, temperature and insolation) had to be assumed to be the same for all villages within the same region. These data shortcomings might have limited the explanatory potential of the environmental variables. Since malaria related and environmental factor data are collected on a continual basis, it would therefore be worthwhile to employ exploratory spatial data analysis techniques to data from subsequent rainy and dry seasons, in order to verify and expand upon current results.

The study advocated the value of geographical spatial analysis for malaria research – an important and topical disease having a crucial impact on the development course of sub-Saharan Africa. Although no stable malaria predictive model could be derived using available data sets, the study provided in a better understanding of the spatial variations in malaria transmission and morbidity, and its spatial interaction with possible environmental determinants. In particular, it provided in a practical demonstration of GIS application capabilities in the arduous search for solutions to this endemic problem. The integration of GIS with statistical methods brought about new insights toward spatial research possibilities, contributing to the important work of elimination and filtering of a range of possible influencing factors – typical of exploratory research. As an initial spatial exploration of the malarial data, this study could therefore be considered as a possible outline for continued and persistent spatial malarial research, in order to restrain the largest killer in Africa.

LIST OF REFERENCES

American Association for the Advancement of Science 1991. *Malaria and Development in Africa - A cross-sectoral approach*. Washington: American Association for the advancement of Science.

Anonymous 1998. Environmental information and health successes in the fight against malaria in Africa. *EIS News* 5:6-9.

Appleton CC, Sharp BL & Le Sueur D 1995. *Wetlands and water-related parasitic diseases of man in southern Africa*. Pretoria, South Africa: Department of Environmental Affairs, South Africa. Pp. 227-246.

Bailey TC & Gatrell AC 1995. *Interactive Spatial Data Analysis*. Harlow, Essex: Longman Scientific & Technical.

Barret A 1986. Medical geography: Concept and definition. In Pacione M (ed.) *Medical Geography: Progress and Prospect*. Pp. 1-34. London: Croom Helm.

Birley MH 1991. *Guidelines for forecasting the vector-borne disease implications of water resource development*. Second Edition. PEEM Guidelines Series 2. Geneva: PEEM secretariat, World Health Organization.

Bradley DJ 1988. The epidemiology of ricefield-associated diseases. In *Vector-borne diseases control in humans through rice agroecosystem management*. Pp. 29-39. Manila: IRRI in collaboration with the WHO/FAO/UNEP PEEM.

Brêtas G 1995. Geographical Information Systems for the study and control of malaria. In De Savigny D & Wijeyaratne P (eds.) *GIS for health and the environment*. Ottawa: International Development Research Centre. Available: www.idrc.ca/books/focus/766/bretas.html [10 June 1997].

Campbell C 1998. Scientists from Africa join forces to push back malaria. *The Star*, 7 April 1998:11.

Capraro I 1997. Genetiese ingryping help dalk in stryd teen dodelike muskiet. *Die Burger*, 6 September 1997:15.

Carnevale P & Robert V 1987. Introduction of irrigation in Burkina Faso and its effects on malaria transmission. In *Effect of agricultural development on vector-borne diseases*. Pp. 57-67. Rome: FAO.

Carter RC, Brook JM & Jewsbury JM 1990. Assessing the impact of small dams on vector borne disease. *Irrigation and Drainage Systems* 4:1-16.

Center for Health Applications of Aerospace Related Technologies (CHAART) (1997, June 6). Landscape Epidemiology and RS/GIS. Earth Science Division, NASA. Available: geo.arc.nasa.gov/esdaff/health/landepi/landepi.html [6 June 1997].

Cliff AD & Haggett P 1988. *Atlas of disease distributions*. Oxford: Basil Blackwell.

Coosemans M & Mouchet J 1990. Consequences of rural development on vectors and their control. *Annales de la Societe Belge de Médecine Tropicale* 70:5-23.

Dossou-Yovo J, Ouattara A, Doannio JMC, Riviere F, Chauvancy G & Meunier JY 1994. Aspect du Paludisme dans un village de Savane humide en Côte d'Ivoire. *Medecine Tropical* 54:331-336.

Douven W & Scholten HJ 1995. Spatial analysis in health research. In De Lepper MJC, Scholten HJ & Stern RM (eds.) *The Added Value of Geographical Information Systems in Public and Environmental Health*. Pp. 117-133. Dordrecht, Netherlands: Kluwer Academic Publishers.

Environmental Systems Research Institute (ESRI) 1992. *Surface modelling with TIN*. Second Edition. Redlands, California: Environmental Systems Research Institute.

Food and Agricultural Organization (FAO) 1984. Pesticide residues in food: report of the joint meeting on pesticide residues. FAO Plant Production and Protection Paper No. 62.

Fuggle RF & Rabie MA (eds.) 1992. *Environmental management in South Africa*. Cape Town: Juta.

Gillet JD 1989. Maintenance and spread of insect-borne disease by the agency of man. In Service M.W. (ed.) *Demography and vector-borne diseases*. Pp. 35-47. Boca, Raton, Florida: CRC Press.

Goudie A (ed.) 1985. *The encyclopaedic dictionary of physical geography*. Oxford: Blackwell.

Gunawardena DM, Muthuwattac L, Weerasingha S, Rajakaruna J, Kumara WU, Senanayaka T, Kotta PK, Wickremasinghe AR, Carter R & Mendis KN 1995. Spatial analysis of malaria risk in an endemic region of Sri Lanka. In De Savigny D & Wijeyaratne P (eds.) *GIS for health and the environment*. Ottawa: International Development Research Centre. Available: www.idrc.ca/books/focus/766/gunawar.html [10 June 1997].

Hay SI, Snow RW & Rogers DJ 1998. Predicting malaria seasons in Kenya using multitemporal meteorological satellite sensor data. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 92:12-20.

Howe GM 1986. Disease mapping. In Pacione M: *Medical Geography: Progress and Prospect*. Pp. 35-63. London: Croom Helm.

International Development Research Centre (IDRC) 1996. Statistics on malaria. Available: www.idrc.ca/books/reports/1996/01-07e.html [10 June 1997].

Jones K & Moon G 1987. *Health, disease and Society: An introduction to medical geography*. London: Routledge & Kegan Paul.

Kim JO & Mueller CW 1978. *Introduction to factor analysis: What it is and how to do it*. Beverly Hills, California: Sage Publications.

Kitron U, Pender 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(5):550-556.

Le Sueur D, Sharp BL & Ngxongo SM 1992. An investigation into the effect of rice cultivation on anopheles mosquito population at Mamfene, with special reference to implications for malaria transmission and control. Office Report. Durban, South Africa: Medical Research Council.

Lichtenberg ER & Getz W 1985. Economics of Rice-field Control in California. *BioScience* 35(5):292-297.

Lindsay SW, Wilkins HA, Zieler HA, Dalv RJ, Petrarca V & Byass P 1991. Ability of *Anopheles gambiae* mosquitoes to transmit malaria during the dry and wet seasons in an area of irrigated rice cultivation in The Gambia. *Journal of Tropical Medicine and Hygiene* 94:313-324.

Löffler H & Malkhazova S 1990. Impacts of wetlands on man. In Patten BC (ed.) *Wetlands and shallow continental water bodies*. Volume 1. Pp. 347-362. The Hague: SPB Academic Publishing.

Loslier L 1995. Geographical Information Systems (GIS) from a Health Perspective. In De Savigny D & Wijeyaratne P (eds.) *GIS for health and the environment*. Ottawa: International Development Research Centre. Available: www.idrc.ca/books/focus/766/loslier1.html [10 June 1997].

MacCormack CP 1984. Human ecology and behavior in malaria control in tropical Africa. *Bulletin of the World Health Organization* 62(Suppl.):81-87.

Martens WJM, Niessen LW, Rotmans J, Jetten TH & McMichael AJ 1995. Potential impact of global climate change on malaria risk. *Environmental Health Perspectives* 103(5):458-464.

Mayer JD 1986. Ecological associative analysis. In Pacione M: *Medical Geography: Progress and Prospect*. Pp. 64-83. London: Croom Helm.

McMillan DE & Meltzer MI 1996. Vector-borne disease control in sub-Saharan Africa: A necessary but partial vision of development. *World Development* 24(3):569-588.

Medical Research Council 1996. MARA – mapping malaria risk in Africa. *MRC News* 27(3):1-3.

Medical Research Council (1997, June 6). Executive summary of the MARA/ARMA project. Available: www.mara.org.za [6 June 1997].

Mulaik SA 1972. *The foundations of factor analysis*. New York: McGraw Hill.

Najera JA 1988. Malaria and rice: strategies for control. In *Vector-borne disease control in humans through rice agroecosystem management*. Pp. 123-132. Manila: IRRI in collaboration with the WHO/FAO/UNEP PEEM.

Ngxongo SM 1993. The epidemiology of malaria in KwaZulu 1980-1991. M.Sc. thesis. Pietermaritzburg, South Africa: University of Natal.

Nobre FF & Carvalho MS 1995. Spatial and Temporal Analysis of Epidemiological Data. In De Savigny D & Wijeyaratne P (eds.) *GIS for health and the environment*. Ottawa: International Development Research Centre. Available: www.idrc.ca/books/focus/766/nobre.html [10 June 1997].

Onori E & Grab B 1980. Indicators for the forecasting of malaria epidemics. *Bulletin of the World Health Organization* 58(1):91-98.

Phillips DR 1990. *Health and health care in the Third World*. New York: Longman Scientific and Technical.

Raybould S, Nicol J, Cross A & Coombes M 1995. The long-term potential of Geographical Information Systems for epidemiology. In De Lepper MJC, Scholten HJ & Stern RM (eds.) *The Added Value of Geographical Information Systems in Public and Environmental Health*. Pp. 277-284. Dordrecht, Netherlands: Kluwer Academic Publishers.

Reuben R 1989. Obstacles to malaria control in India – the human factor. In Service MW (ed.) *Demography and vector-borne diseases*. Pp. 143-154. Boca, Raton, Florida: CRC Press.

Robert V, Gazin P, Boudin C, Molez JF, Ouedraogo V & Carnevale P 1985. La transmission du paludisme en zone de savanne arborée et en zone rizicole des environs de Bobo-Dioulasso (Burkina Faso). *Annales de la Societe Belge de Médecine Tropicale* 65(Suppl. 2):201-214.

Roger PA & Bhuiyan SI 1990. Ricefield ecosystem management and its impact on disease vectors. *International Journal of Water Resources Development* 6(1):2-18.

Scholten HJ & De Lepper MJC 1991. The benefits of the application of Geographical Information Systems in public and environmental health. *WHO Statistical Quarterly* 44(3).

Service MW 1989. Irrigation: boon or bane? In Service M (ed.) *Demography and Vector-borne Diseases*. Pp. 237-254. Boca, Raton, Florida: CRC Press.

Service MW 1993. *Mosquito Ecology: Field sampling methods*. Second Edition. London: Elsevier Applied Science.

Sharp BL 1991. Aspects of the epidemiology of malaria in Natal. PhD dissertation. Durban, South Africa: University of Natal Medical School.

Sharp BL, Quicke FC & Jansen EJ 1984. Aspects of the behaviour of five anopheline species in the endemic malaria area of Natal. *Journal Entomological Society of Southern Africa* 47:251-258.

Smith T, Hurt N, Teuscher T & Tanner M 1995. Is fever a good sign for clinical malaria in surveys of endemic communities? *American Journal of Tropical Medicine and Hygiene* 52(4):306-310.

Snow RW, Bastos de Azevedo I, Lowe BS, Kabiru EW, Nevill CG, Mwankusye S, Kassiga G, Marsh K & Teuscher T 1994. Severe childhood malaria in two areas of markedly different *falciparum* transmission in East Africa. *Acta Tropica* 57:289-300.

Snow RW & Marsh K 1995. Will reducing *Plasmodium falciparum* transmission alter malaria mortality among African children? *Parasitology Today* 11(5):188-194.

Snow RW, Marsh K & Le Sueur D 1996. The need for maps of transmission intensity to guide malaria control in Africa. *Parasitology Today* 12(12):455-456.

Snow RW, Omumbo JA, Lowe B, Molyneux CS, Obiero JO, Palmer A, Weber MW, Pinder M, Nahlen B, Obonyo C, Newbold C, Gupta S & Marsh K 1997. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *The Lancet* 349:1650-1654.

Snow WF 1983. Mosquito production and species succession from an area of irrigated ricefields in The Gambia, West Africa. *Journal of Tropical Medicine and Hygiene* 86:237-245.

Stadler C 1998. A matter of life or death. *Geoinformation Africa* 5:20-24.

Stamp D & Clark AN (eds.) 1979. *A glossary of geographical terms: based on a list / prepared by a Committee of the British Association for the Advancement of Science*. Third Edition. London: Longman.

Stern RM 1995. Environment and health data in Europe as a tool for risk management: needs, uses and strategies. In De Lepper MJC, Scholten HJ & Stern RM (eds.) *The Added Value of Geographical Information Systems in Public and Environmental Health*. Pp. 3-24. Dordrecht, Netherlands: Kluwer Academic Publishers.

Strebel PM & Thompson ML 1990. Time series modelling of malaria notification data. *The Southern African Journal of Epidemiology and Infection* 5(2):27-32.

Stuttaford MC 1994. Aspects of a GIS for medical geographers and malaria control. M.Sc. Thesis. University of Natal, South Africa.

Teuscher T 1996. Association between Rice Production Systems and Vector-borne Diseases in West Africa: Technical and financial report, activities for the period May 1995 – April 1996. Office Report. Bouaké, West Africa Rice Development Association (WARDA).

Teuscher T 1999. Information given to CM van Wijk during a personal communication in 1999. Public Health Epidemiologist, Health Research Consortium, WARDA, Bouaké, Côte d'Ivoire.

Tran DV, Ofori C & Berney O 1995. Inland swamp rice-based development in Africa. *International Rice Commission Newsletter* 44:7-20.

Trape JF & Rogier C 1996. Combatting malaria morbidity and mortality in tropical Africa by reducing transmission: a battle lost from the start? *Parasitology Today* 12(6):236-240.

Van de Giesen N 1997. Information given to CM van Wijk during a personal communication in 1997. Hydrologist, WARDA, Bouaké, Côte d'Ivoire.

Verhoef H & Bos R 1992. Development and vector-borne diseases. *World Health* Nov-Dec 1992:15-17.

White G 1976. The main effects and problems of irrigation. *Arid land irrigation in developing countries*. International Symposium. February. Alexandria, Egypt. Pp. 16-21.

Wood BL, Beck LR, Washino RK, Hibbard KA & Salute JS 1992. Estimating high mosquito-producing rice fields using spectral and spatial data. *International Journal of Remote Sensing* 13(15):2813-2826.

World Bank 1994. *Better health in Africa: Experience and lessons learned*. Washington: World Bank.

World Health Organization (WHO) 1980. *Environmental Management for vector control*. Geneva: World Health Organization, Vector Biology and Control Division.

World Health Organization (WHO) 1992. World malaria situation in 1990. *World Health Statistical Quarterly* 45:257-266.

Yoon SS 1995. Geographical Information Systems: A new tool in the fight against Schistosomiasis. In De Lepper MJC, Scholten HJ & Stern RM (eds.) *The Added Value of Geographical Information Systems in Public and Environmental Health*. Pp. 201-213. Dordrecht, Netherlands: Kluwer Academic Publishers.

APPENDIX A: FACTOR LOADINGS**Table A1: Factor Loadings**

Factor Name	FACTOR_1 (Humidity)	FACTOR_2 (Water body size)	FACTOR_3 (Temperature)	FACTOR_4 (Distance to water bodies)	FACTOR_5 (Barrenness)
SUPCUS	-0.081603	0.837453	0.130274	-0.035670	0.023979
SUPERFIC	-0.069685	0.819636	0.101983	0.055714	0.241406
NEARWETD	0.056026	0.077288	-0.146401	0.832799	-0.267895
TEMPMAX	-0.792080	-0.038769	0.572294	0.040460	0.009047
TEMPMIN	0.557594	0.189181	0.670763	-0.122342	0.289091
TEMPMOY	-0.160293	0.052412	0.985962	-0.012159	0.061395
HUMIDMAX	0.935598	-0.005372	0.140789	0.023923	-0.094143
HUMIDMIN	0.975480	-0.038700	0.146266	0.035684	-0.119073
HUMIDMOY	0.974907	-0.021254	0.146564	0.030005	-0.107933
VILWETDI	-0.010525	-0.389126	0.090891	0.743562	0.331674
SUMRAIN	0.865974	0.071338	-0.290571	-0.067100	0.131183
RAINDAYS	0.861237	0.167762	-0.286217	-0.009336	0.105010
MEANFLOO	0.089257	0.925543	0.016892	0.015575	0.131596
FLOODDAY	0.125588	0.675450	-0.133884	-0.518226	-0.162144
WFLOODPE	0.211799	0.832131	-0.078874	-0.234001	0.087808
VEGETATI	0.071193	-0.412282	-0.151495	0.034753	-0.766207
Expl. Var	5.321415	3.781054	2.083062	1.599669	1.029372
Prp. Totl	0.332588	0.236316	0.130191	0.099979	0.064336

APPENDIX B: SEMIVARIOGRAMS

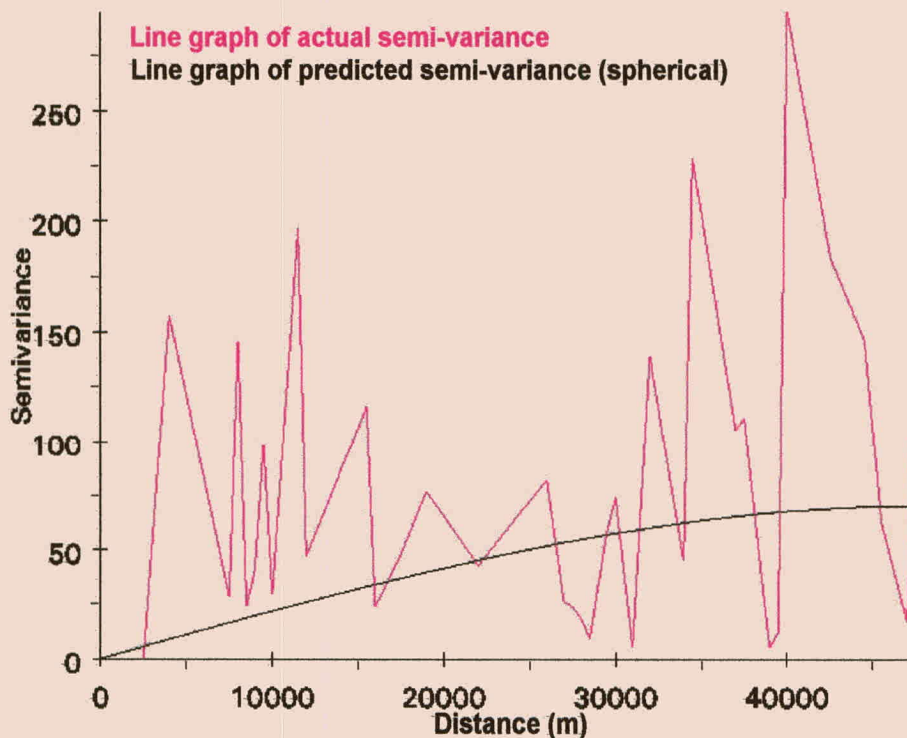


Figure B1: Semivariogram of fever episodes, Survey 1, and its spherical estimation

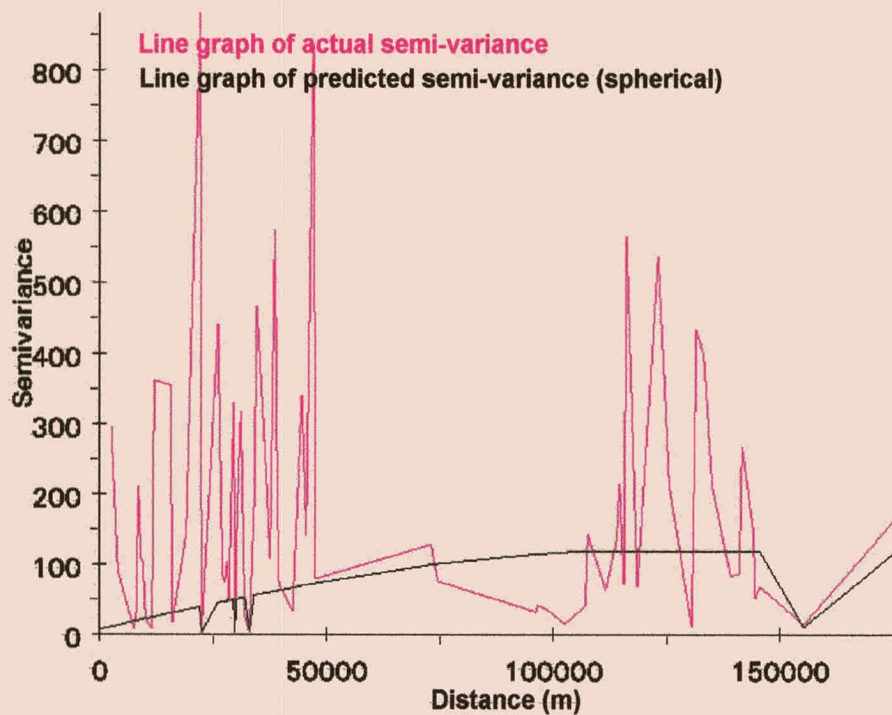


Figure B2: Semivariogram of fever episodes, Survey 2, and its spherical estimation

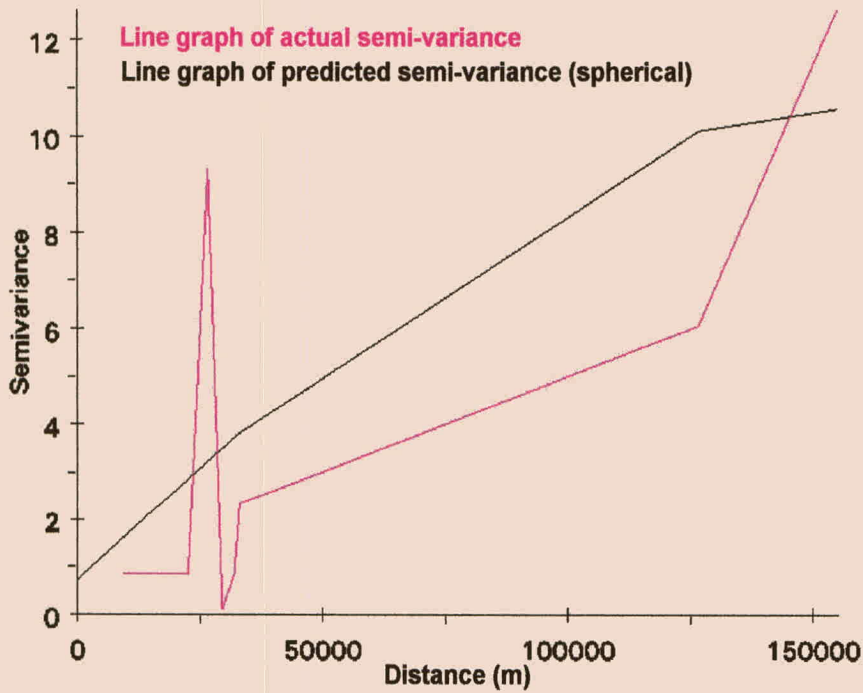


Figure B3: Semivariogram of *An. funestus* transmission intensity, Survey 1, and its spherical estimation

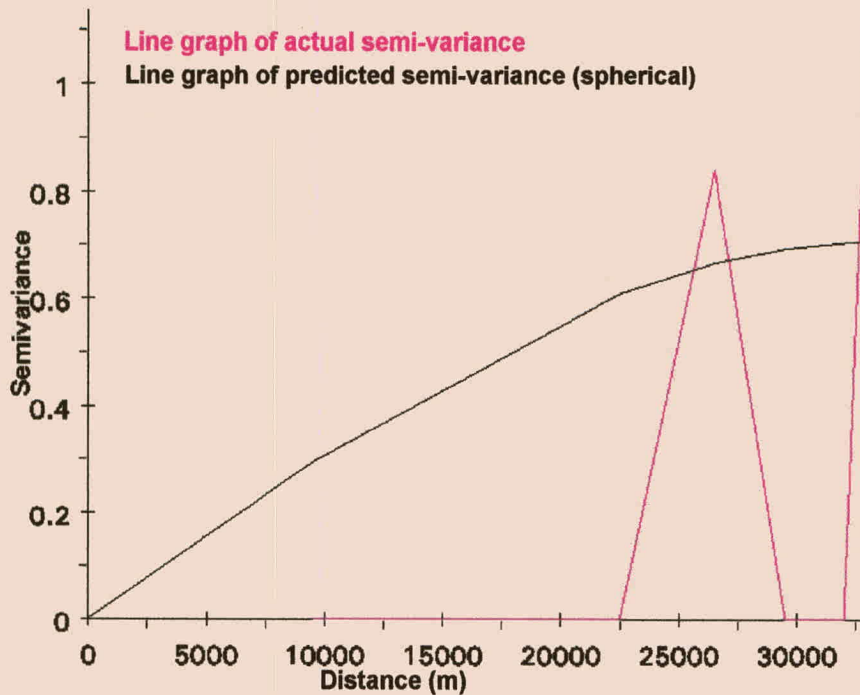


Figure B4: Semivariogram of *An. funestus* transmission intensity, Survey 2, and its spherical estimation

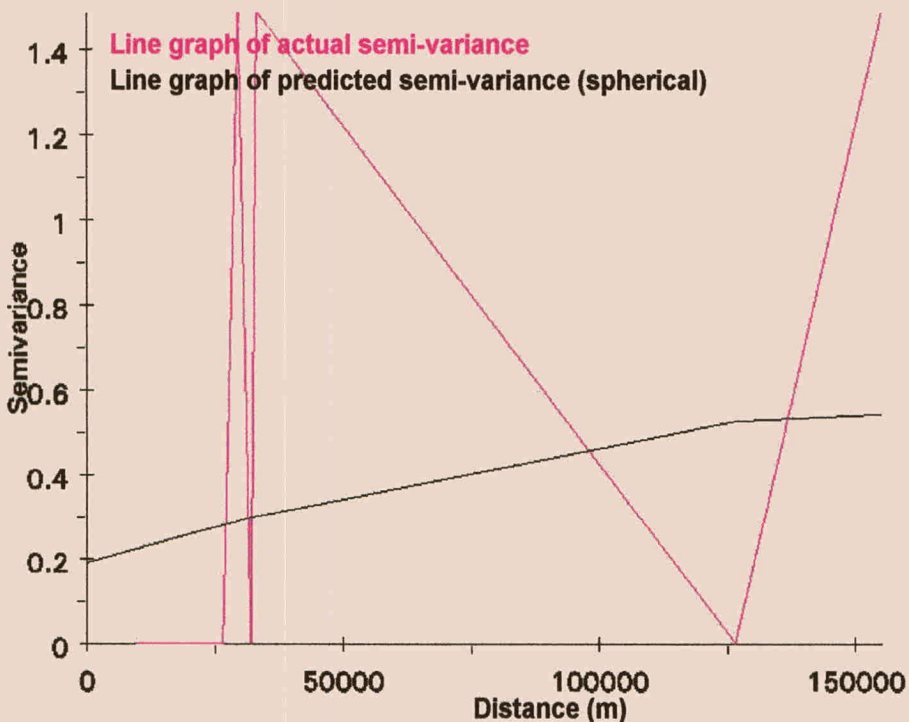


Figure B5: Semivariogram of *An. funestus* transmission intensity, Survey 3, and its spherical estimation

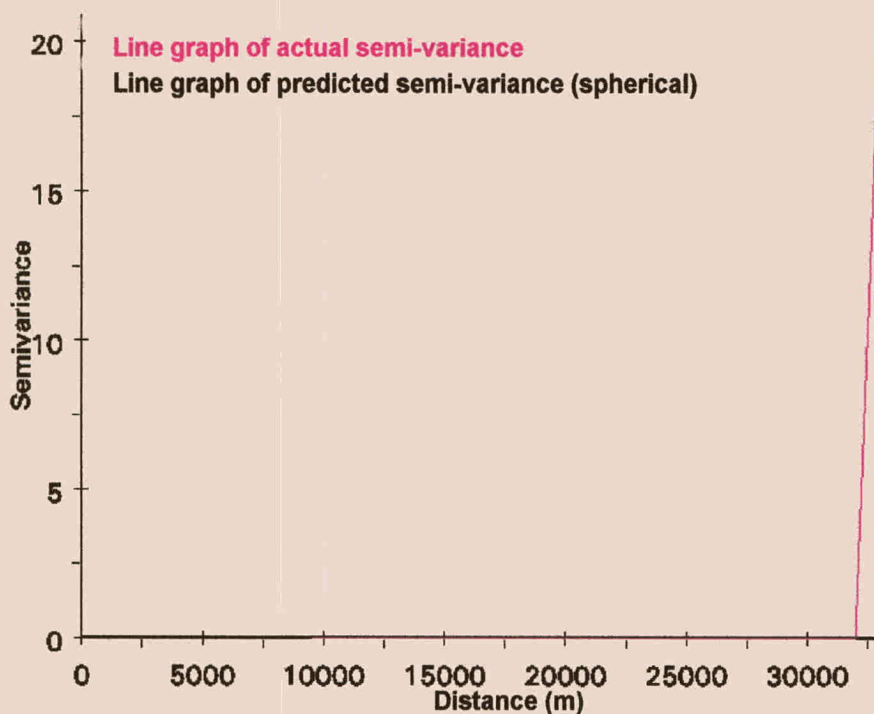


Figure B6: Semivariogram of *An. funestus* transmission intensity, Survey 4, and its spherical estimation

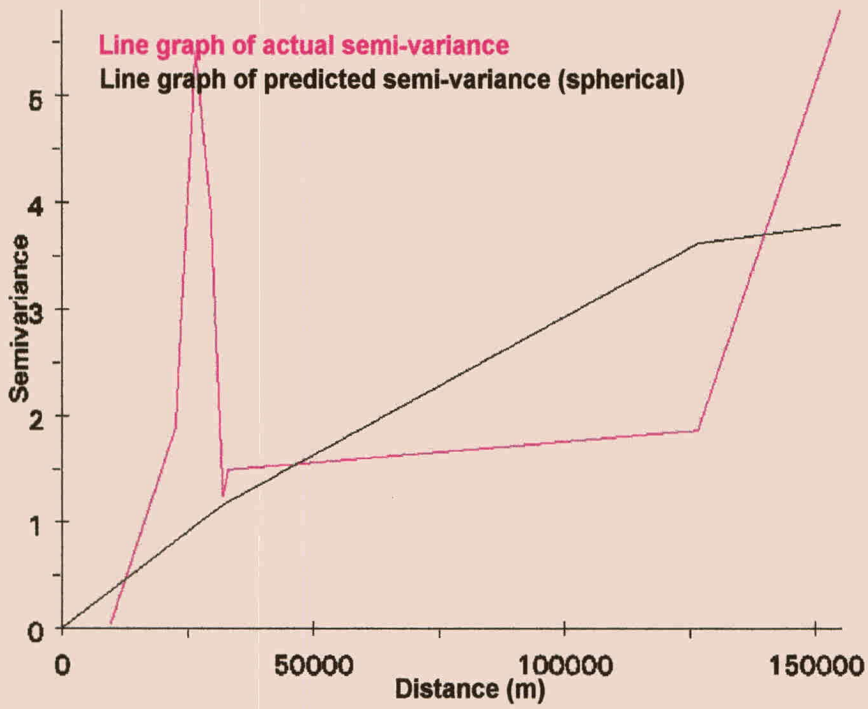


Figure B7: Semivariogram of *An. gambiae* transmission intensity, Survey 1, and its spherical estimation

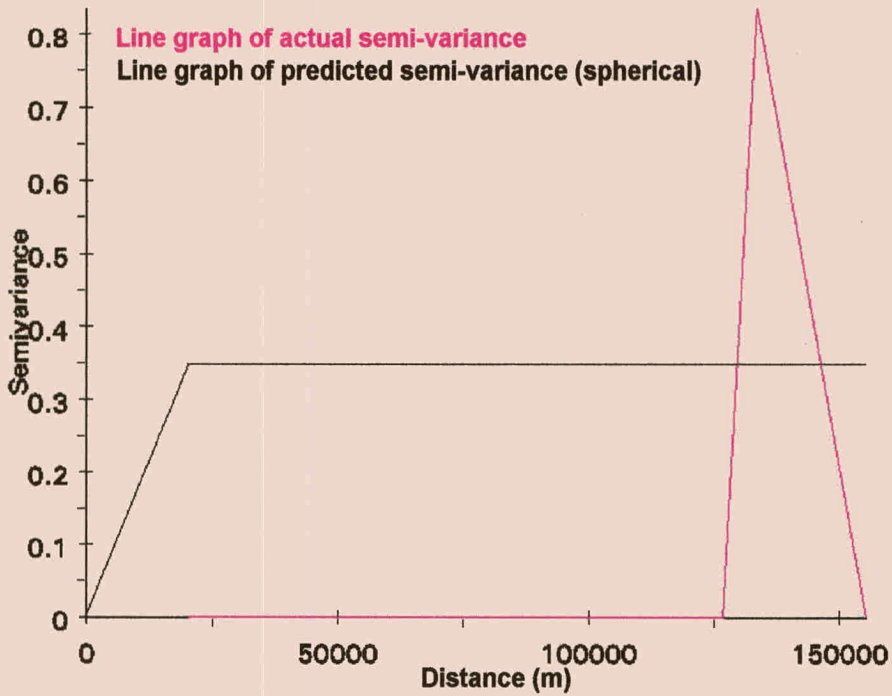


Figure B8: Semivariogram of *An. gambiae* transmission intensity, Survey 2, and its spherical estimation

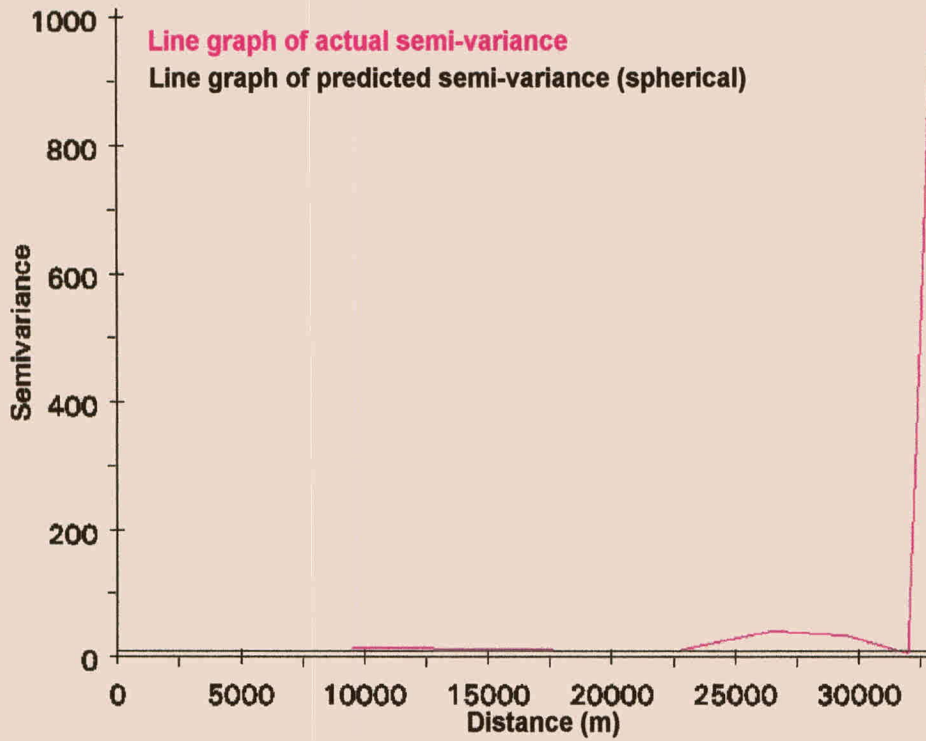


Figure B9: Semivariogram of *An. gambiae* transmission intensity, Survey 3, and its spherical estimation

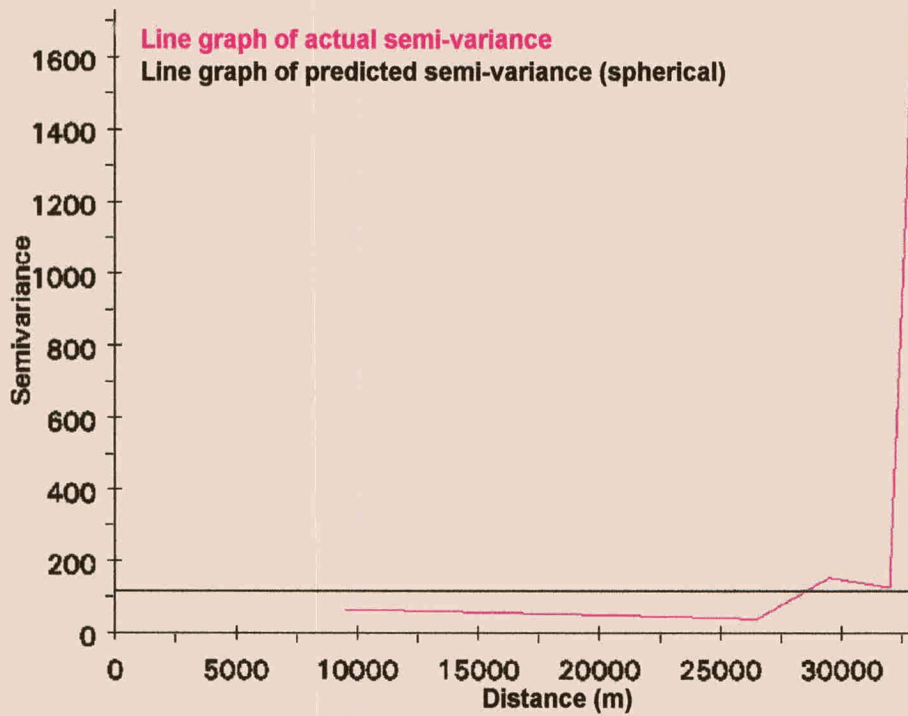


Figure B10: Semivariogram of *An. gambiae* transmission intensity, Survey 4, and its spherical estimation