Towards malaria prediction in Sri Lanka: Modelling spatial and temporal variability of malaria case counts

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät der

Universität Basel

von

Olivier J.T. Briët

aus

den Niederlanden

Basel, 2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel auf Antrag der Herren und Frauen Prof. Dr. Marcel Tanner, PD Dr. Penelope Vounatsou, Dr. Immo Kleinschmidt, Prof. Dr. Thomas Smith.

Basel, den 24. Juni 2008

Prof. Dr. Hans-Peter Hauri

Dekan der Philosophisch-Naturwissenschaftlichen Fakultät

Table of contents

Acknowledgments

Summary

This thesis was motivated by the need of the Anti Malaria Campaign (AMC) of Sri Lanka for malaria risk maps and malaria case number predictions to assist in the planning for malaria control. Despite a wealth of high resolution data collected over decades, a malaria forecasting system was not in place, and detailed island-wide maps of malaria incidence could permit the assessment of the malaria situation and its determinants. The overall aim of this thesis was to describe the spatial and seasonal distribution of malaria in Sri Lanka and associated factors, and to develop a malaria forecasting system.

In this thesis, the spatial variation of malaria in Sri Lanka was described in relation to risk factors. Also, the risk and the impact of a tsunami natural disaster on malaria transmission and malaria control in Sri Lanka were evaluated. The relation in space between seasonality of malaria and seasonality of rainfall, and the relationship between monthly malaria case time series and monthly rainfall time series in Sri Lanka were quantified. A model for short term malaria prediction was developed and implemented in Sri Lanka for use by the AMC. This thesis also contributed a statistical methodology for analysing over dispersed temporal count data with non stationary and / or seasonal behaviour, such as observed in malaria case count time series in Sri Lanka.

In Chapter 1, the stage was set by briefly describing malarial disease and the biology of malarial parasites and vectors relevant to Sri Lanka. The influence of weather on malaria transmission, and observed linkages between weather and malaria in terms of spatial and temporal patterns were introduced. Immunity was also briefly discussed, because it affects the translation of (unobserved) disease transmission patterns into patterns of observed malaria cases. A brief overview was given of the history of malaria and malaria control in Sri Lanka.

Chapter 2 provided health professionals and the larger general public with the first island-wide incidence maps of *Plasmodium vivax* and *Plasmodium falciparum* malaria at sub district resolution. The distribution and seasonality of *P. vivax* and *P. falciparum* incidence was remarkably similar within each district, although they varied spatially. The annual malaria incidence changed over the 1995 – 2002 period, and the rate of change varied with the area, thus indicating the need for regular

updates of the incidence maps. The spatial and temporal malaria distribution in the country was related to accessibility of areas for implementation of malaria control (in particular governed by the armed conflict and the peace process), and to socio economic and environmental factors. Also, the exposure of tourists to malaria infection was discussed.

Chapter 3 provided a re-assessment of the malaria situation, including details on vector insecticide resistance, parasite drug resistance, and insights into the national policy for malaria diagnosis and treatment. The assessment and its publication were triggered by the tsunami that hit on 26 December 2004, and the ensuing international concern about possibilities of an increase of vector borne diseases. The likelihood of a widespread outbreak was estimated as limited. The public health system was deemed capable of dealing with the possible threat of a malaria outbreak. Concerns were expressed that the influx of foreign medical assistance, drugs, and insecticides could interfere with malaria surveillance, and the long term malaria control strategy of Sri Lanka, if not in accordance with government policy.

Chapter 4 assessed the impact of the tsunami on the malaria situation and the national and international malaria control efforts in the year following the tsunami. Malaria incidence had decreased in most districts, including the ones that were hit hardest by the tsunami, and the whole-country malaria incidence time series did not deviate from the downward trend that started in 2000. The focus of national and international post tsunami malaria control efforts was supply of antimalarials, distribution of impregnated mosquito nets and increased monitoring in the affected area. Internationally donated antimalarials were either redundant or did not comply with national drug policy. There was no indication of increased malaria vector density.

In Chapter 5, the spatial correlation between average seasonality of malaria and climatic seasonality of rainfall was studied. A simple index for seasonality was developed by making use of the characteristic of a varying degree of bimodality of seasonality present in both malaria and rainfall in Sri Lanka. The malaria seasonality index was significantly associated with the rainfall seasonality index in a regression taking spatial autocorrelation into account. This was in paradox with the negative correlation in space between annual rainfall and malaria endemicity (Chapter 2). Both rainfall and malaria may react independently to monsoonal periodicity, but given the fact that rainfall has an important impact on the availability and quality of breeding

sites for malaria vectors, it is clear that rainfall seasonality is an important driver of malaria seasonality.

In Chapter 6, the temporal correlation between monthly malaria case time series and monthly rainfall time series was explored for each district separately. For most districts, strong positive correlations were found for malaria time series lagging zero to three months behind rainfall. However, only for a few districts, weak positive (at lags zero and one) or weak negative (at lags two to six) correlations were found if autocorrelation and seasonality were removed from the series prior to cross-correlation analysis, thus indicating that rainfall might have little potential use in a malaria forecasting system. These cross correlation analyses had the drawbacks that inter-annual effects were masked due to detrending of the data, and that potentially seasonally varying effects were not taken into account. Subsequent inter-annual analysis showed strong negative correlations between malaria and rainfall for a group of districts in the centre-west of the country. Seasonal inter-annual analysis showed that the effect of rainfall on malaria varied according to the season (and geography).

Chapter 7 focused on the development of a malaria forecasting system for Sri Lanka, which could assist in the efficient allocation of resources for malaria control, especially when malaria is unstable and fluctuates in intensity both spatially and temporally. Several types of time series models were tested in their ability to predict the monthly number of malaria cases in districts one to four months ahead. Different districts required different prediction models, and the prediction accuracy varied with district and forecasting horizon. It was subsequently tested if rainfall or malaria patterns in neighbouring districts could improve prediction accuracy of the selected models. Only for a few districts, a modest improvement was made when rainfall was included in the models as a covariate. This modest improvement was not deemed sufficient to merit investing in a forecasting system for which rainfall data are routinely processed. The development and launch of a system for forecasting malaria by the AMC was described in addendum to Chapter 7.

Throughout the statistical modelling in Chapter 7, it was assumed that logarithmically transformed malaria case data were approximately Gaussian distributed. However, such an approximation is less close when case numbers are low, as was the case at the time of writing. Therefore, in Chapter 8, a class of generalised multiplicative seasonal autoregressive integrated moving average models for the parsimonious and

observation-driven modelling of non Gaussian, non stationary and / or seasonal time series data was developed.

Chapter 9 provides a general discussion in which the contributions of this thesis are put into context, in which limitations of this thesis are discussed and directions for future research outlined.

Zusammenfassung

Diese These wurde durch die Notwendigkeit von Malaria-Risiko-Karten und Malaria Fall Zahl Vorhersagen anhand der Anti-Malaria-Kampagne (AMC) von Sri Lanka zur Unterstützung bei der Planung für die Malariabekämpfung motiviert. Trotz einer Fülle von hochauflösenden Daten über Jahrzehnte, gab es keine Vorhersage-Systeme für Malaria in Sri Lanka und es gab keine detaillierten Inselkarten von Malariainzidenz für die Bewertung der Malariasituation und ihrer Determinanten. Das übergeordnete Ziel dieser Arbeit war die Beschreibung von der räumlichen und saisonalen Verteilung von Malaria in Sri Lanka und den damit verbundenen Faktoren sowie die Entwicklung eines Malaria-Vorhersage-System.

In dieser Arbeit wurde die räumliche Variation der Malaria in Sri Lanka in Bezug auf die Risikofaktoren beschrieben. Auch das Risiko und die Auswirkungen der Tsunami-Katastrophe auf die natürliche Übertragung der Malaria und Malaria Bekämpfung in Sri Lanka wurde ausgewertet. Die räumliche Beziehung zwischen Malariasaisonabhängigkeit und klimatische Saisonalität der Niederschläge, sowie die Beziehung zwischen den monatlichen Malariafallzeitreihen und monatlichen Niederschlagsmengezeitreihen in Sri Lanka wurden quantifiziert. Ein Modell für die kurzfristige Malaria Vorhersage wurde entwickelt und für die Verwendung durch den AMC in Sri Lanka implementiert. Diese Arbeit trug auch eine statistische Methode für die Analyse zeitlich überverteilter Daten mit nicht stationären und / oder saisonalem Verhalten, wie Malariafallzeitreihen in Sri Lanka, bei.

Kapitel 1 gab eine kurze Beschreibung der Malariaerkrankung und der Biologie der Malariaparasiten und Vektoren in Sri Lanka. Der Einfluss der Witterung auf die Übertragung der Malaria und beobachtete Zusammenhänge zwischen Witterung und Malaria im Hinblick auf die räumlichen und zeitlichen Muster wurden eingeführt. Immunität wurde auch kurz diskutiert, denn es wirkt sich auf die Umwandlung von (unbeobachteten) Übertragungsmustern der Krankheit in Muster von beobachteten Malariafällen aus. Eine kurze Übersicht von der Geschichte der Malaria und Malaria Bekämpfung in Sri Lanka wurde gegeben.

Kapitel 2 lieferte für die Fachkräfte des Gesundheitswesens und der weiten Öffentlichkeit die ersten inselweiten Inzidenzkarten von *Plasmodium vivax* und *Plasmodium falciparum* Malaria mit Auflösung bis zu den Unterbezirken . Die

Verteilung und Seasonalität der *P. vivax*- und *P. falciparum*-Inzidenz war bemerkenswert ähnlich in jedem Bezirk, auch wenn sie räumlich variierten. Die jährliche Malariainzidenz änderte sich von 1995 – 2002 und das Tempo der Veränderung variierte mit der Gegend. So zeigte sich die Notwendigkeit einer regelmäßigen Erneuerung der Inzidenzkarten. Die räumliche und zeitliche Verteilung der Malaria im Land war im Zusammenhang mit der Zugänglichkeit von Gegenden zur Malariabekämpfung (insbesondere durch den bewaffneten Konflikt und den Friedensprozess), und die sozioökonomischen und ökologischen Faktoren. Auch die Exposition von Touristen zu Malariainfektion wurde diskutiert.

Kapitel 3 verschaffte eine erneute Bewertung der Malariasituation, einschließlich Details über Vektorinsektizidresistenz, Parasitenmedikamentenresistenz und Einblicken in die nationale Politik für die Diagnose und Behandlung der Malaria. Die Bewertung und die Veröffentlichung wurden ausgelöst durch den Tsunami, vom 26. Dezember 2004, und die internationale Besorgnis über die sich daraus ergebende mögliche Erhöhung der Vektorkrankheiten. Die Wahrscheinlichkeit eines weit verbreiteten Ausbruchs wurde als begrenzt geschätzt. Das System der öffentlichen Gesundheit wurde als geeignet für den Umgang mit der möglichen Gefahr einer Malariaepidemie erachtet. Bedenken wurden geäußert, dass der Zustrom von ausländischer medizinischer Hilfe, Drogen und Insektiziden die Malariaüberwachung und die langfristige Kontrolle der Malariastrategie von Sri Lanka stören könnte, wenn sie nicht im Einklang mit der Regierungspolitik waren.

Kapitel 4 bewertet die Auswirkungen des Tsunami auf die Malariasituation und die nationalen und internationalen Anstrengungen in der Kontrolle der Malaria im Jahr nach dem Tsunami. Malariainzidenz war gesunken in den meisten Bezirken, darunter diejenigen, die am härtesten von der Tsunami-Katastrophe getroffen waren, und die Malariainzidenz des ganzen Landes weichte nicht vom Abwärtstrend ab der im Jahr 2000 begann. Der Fokus der nationalen und internationalen Post-Tsunami Malariakontrollbemühungen bestand in der Versorgung mit Malariamitteln, Verteilung von imprägnierten Moskitonetzen und eine stärkere Überwachung in den betroffenen Bereichen. International gespendete Malariamittel wurden entweder überflüssig oder nicht im Einklang mit der nationalen Drogenpolitik. Es gab keine Hinweise auf eine erhöhte Malaria-Vektor-Dichte.

In Kapitel 5 wurde die räumliche Korrelation zwischen dem Durchschnitt der Saisonabhängigkeit von Malaria und klimatische Saisonalität der Niederschläge untersucht. Ein einfacher Index für saisonal wurde entwickelt durch die Verwendung des Merkmals des unterschiedlichen Grades der Zweigipfligkeit der Saisonalität der Malaria und der Niederschläge in Sri Lanka. Der Malariasaisonalitätsindex wurde signifikant mit dem Niederschlagssaisonalitätsindex in einer Regression assoziiert, die räumliche Autokorrelation berücksichtigte. Das war im Gegensatz zur negativen Korrelation im Raum zwischen den jährlichen Niederschlägen und der endemischen Situation von Malaria (Kapitel 2). Sowohl Niederschläge als auch Malaria könnten unabhängig auf die monsunische Periodizität reagieren, aber angesichts der Tatsache, dass die Niederschläge einen wichtigen Einfluss auf die Verfügbarkeit und Qualität der Brutplätze für die Malariavektoren haben, ist es wahrscheinlich, dass Niederschlagssaisonalität ein wichtiger Faktor für Malariasaisonabhängigkeit ist.

In Kapitel 6, wurde die zeitliche Korrelation zwischen monatlichen Malariafallzeitreihen und monatlichen Niederschlagsmengezeitreihen für jeden Bezirk getrennt erforscht. Für die meisten Distrikte, wurde für Malariazeitreihen eine starke positive Korrelation gefunden für den Zeitraum von null bis drei Monaten nach Niederschlägen. Aber nur für ein paar Bezirke mit schwach positiven (Verzögerungen bei null und einem Monat) oder schwach negativen (bei Verzögerungen von zwei bis sechs Monate) wurde Korrelationen gefunden, wenn Autokorrelation und Saisonalität aus der Serie vor der Cross-Korrelation Analyse herausgelöst wurden, was bedeutet, dass Niederschläge vielleicht wenig Potenzial in der Nutzung eines Malaria-Vorhersage-System hat. Diese Cross-Korrelation Analysen hatten die Nachteile, dass zwischenjährlichen Auswirkungen verdeckt waren durch Trendherausnahme der Daten und mögliche saisonal unterschiedliche Effekte wurden nicht berücksichtigt. Die nachfolgende zwischenjährliche Analyse zeigte starke negative Korrelationen zwischen Malaria und Niederschlägen für eine Gruppe von Bezirken im mittleren Westen des Landes. Eine saisonale zwischenjährliche Analyse zeigte, dass die Wirkung von Regen auf Malaria je nach der Saison (und Geographie) differierte.

Kapitel 7 konzentriert sich auf die Entwicklung eines Malariavorhersagesystems für Sri Lanka, dass bei der effizienten Bereitstellung von Ressourcen für die Kontrolle der Malaria förderlich sein könnte, vor allem, wenn die Malaria instabil ist und sowohl räumlich und zeitlich in der Intensität schwankt. Mehrere Arten von

Zeitreihenmodelle wurden getestet und ihre Fähigkeit zur Vorhersage der monatlichen Zahl der Malariafälle in den Bezirken ein bis vier Monate im Voraus. Verschiedene Bezirke benötigten verschiedene Vorhersage-Modelle, und die Genauigkeit der Vorhersage änderte sich mit Bezirks- und Prognosehorizont. Es wurde dann getestet, wie Regen- oder Malariamuster in den benachbarten Bezirken die Genauigkeit der Vorhersage der ausgewählten Modelle verbessern könnte. Nur für ein paar Bezirke wurde mit Niederschlägen als Variable in den Modellen eine bescheidene Verbesserung erhalten. Diese leichte Verbesserung war nicht als ausreichend erachtet, um eine Investition in ein Vorhersagesystem, bei dem Niederschlagsdaten routinemäßig verarbeitet werden müssen, zu begründen. Die Entwicklung und Einführung eines Systems für die Vorhersage von Malaria durch die AMC wurde in Ergänzung zu Kapitel 7 beschrieben.

Während der statistischen Modellierung in Kapitel 7 war als gegeben vorausgesetzt, dass logarithmisch transformierte Malariafalldaten ungefähr normal verteilt waren. Allerdings, eine solche Annäherung ist weniger gut, wenn Fallzahlen niedrig sind, wie zum Zeitpunkt der Redaktion der Fall war. Daher wurde in Kapitel 8 eine Klasse von allgemeinen multiplikativen saisonalen integrierten autoregressiven Modellen mit gleitendem Durchschnitten für die Beobachtung und sparsame Modellierung von nicht gaußförmigen, nicht stationären und / oder saisonalen Zeitreihendaten entwickelt.

Kapitel 9 enthält eine allgemeine Diskussion, in denen die Beiträge dieser Arbeit in Kontext gebracht worden sind, in denen die Grenzen dieser Arbeit diskutiert worden sind, und in denen die Richtungen für die zukünftige Forschung skizziert worden sind.

Abbreviations

ACF	Auto Correlation Function
ACP	Autoregressive Conditional Poisson
ADRA	Adventist Development and Relief Agency International
AIC	Akaike's Information Criterion
AMC	Anti Malaria Campaign
API	Annual Parasite Incidence
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
CAR	Conditional AutoRegressive
CD-ROM	Compact Disc - Read Only Memory
CLAR	Conditional Linear AutoRegressive
CWS	Church World Service
DDT	DichloroDiphenylTrichloroethane
DIC	Deviance Information Criterion
DNA	DeoxyriboNucleic Acid
DPDH	Deputy Provincial Directors of Health
DSD	Divisional Secretariat Division
EIR	Entomological Inoculation Rate
ENSO	El Niño Southern Oscillation
EPA	Environmental Protection Agency (United States)
EPRI	Electric Power Research Institute (United States)
GARIMA	Generalised AutoRegressive Integrated Moving Average
GARMA	Generalised AutoRegressive Moving Average
GFATM	Global Fund to fight Aids, Tuberculosis and Malaria
GIS	Geographical Information Systems
GNU	GNU's Not Unix
GSARIMA	Generalised multiplicative Seasonal AutoRegressive Integrated Moving Average
ICRC	International Committee of the Red Cross
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
IDP	Internally Displaced Persons
INAR	INteger-valued AutoRegressive

INARMA	INteger-valued AutoRegressive Moving Average
INMA	INteger-valued Moving Average
IOM	International Organization for Migration
IRS	Indoor Residual Spraying
IWMI	International Water Management Institute
LLIN	Long Lasting Insecticide treated mosquito Net
LWR	Lutheran World Relief
MCMC	Markov Chain Monte Carlo
MIM	Multilateral Initiative on Malaria
MOH	Medical Officer of Health
NA	Not Available
NDVI	Normalized Difference Vegetation Index
NGO	Non-Governmental Organizations
NOAA	National Oceanic and Atmospheric Administration (United States)
NSF	National Science Foundation (United States)
OCHA	United Nations Office for the Coordination of Humanitarian Affairs
OPD	Out Patient Department
PACF	Partial Auto Correlation Function
PAR	Poisson AutoRegressive
PQ	Primaquine and Chloroquine
Q-Q	Quantile-Quantile
RBM	Roll Back Malaria
RMO	Regional Malaria Officer
RSS	Residual Sum of Squares
SARIMA	Seasonally mixed Auto Regressive Integrated Moving Average
SARMA	Seasonal AutoRegressive Integrated Moving Average
SDC	Swiss agency for Development and Cooperation
SIMA	System wide Initiative on Malaria and Agriculture
SLRCS	Sri Lankan Red Cross
SP	Sulphadoxine-Pyrimethamine
TEDHA	Tropical and Environmental Diseases and Health Association
UN	United Nations
UNHCR	United Nations High Commissioner for Refugees
UNICEF	United Nations Children's Fund

WHO   World Health Organization

WinBUGS   Bayesian inference Using Gibbs Sampling for Windows

List of tables

List of figures

Chapter 1 – General Introduction

Malaria

Malaria is a disease caused by an infection of a eukaryote parasite of the genus *Plasmodium* (Haemosporida: Plasmodiidae) of the phylum Apicomplexa. Of the four species of *Plasmodium* that infect humans (*P. falciparum*, *P. malariae, P. ovale*, and *P. vivax*), *P. falciparum* and *P. vivax* are currently the only species reported in Sri Lanka.

Plasmodium falciparum infection can be fatal when it develops into a severe malaria, and most of the fatal malaria cases in the world are attributed to this species. Red blood cells infected with *P. falciparum* can clog small blood vessels in the brain, resulting in cerebral malaria, which is often fatal. Also, *P. falciparum* is often associated with severe anaemia because it multiples rapidly in the blood. Other nonspecific symptoms include chills, discomfort, fatigue, headache, muscle pain, cough and respiratory distress, low birth weights, hypoglycaemia and organ failure.

Plasmodium vivax can cause death due to enlarged spleen, but more often it causes debilitating, but not deadly, nonspecific symptoms as listed above.

Plasmodium falciparum and *P. vivax* have a similar biology, but there are important differences. For example, *P. vivax* has dormant liver stages termed "hypnozoites". Also, *P. vivax* is not associated with cerebral malaria.

Lifecycle of *Plasmodium*

The plasmodia, in the sporozoite stage, enter the bloodstream of the vertebrate host when excreted from the salivary gland of an infected female mosquito during the process of biting (only female mosquitoes take blood meals). Because of blood circulation, the sporozoites are carried to the liver within minutes. Each sporozoite then invades a liver cell (hepatocyte). A trait of *e. g. P. vivax* is that some of the sporozoites do not immediately start to grow and divide (schizogony) after entering the hepatocyte, but remain in a dormant, hypnozoite stage for weeks, months, or even years [1]. Thus a single infection can be responsible for a series of "relapses". The factors that eventually trigger schizogony are not known. The exoerythrocytic schizogony stage is characterized by multiple rounds of nuclear division without cellular segmentation, which takes about six to fifteen days, depending on the *Plasmodium* species. After a number of nuclear divisions, the schizont will segment and merozoites are formed. After rupturing of the schizont, the merozoites are

released in the bloodstream. Each merozoite invades a red blood cell (erythrocyte) and transforms into a trophozoite, a feeding stage. After growing, the trophozoite transforms itself into a schizont by replicating its DNA multiple times without cellular segmentation. The schizont then undergoes cellular segmentation and differentiation to form a number of merozoite cells in the erythrocyte. The merozoites burst from the erythrocyte, and enter other erythrocytes within minutes. The rupture of the membrane of the parasitized erythrocytes is associated with fever in the host. Some trophozoites do not develop into schizonts, but develop into a micro or a macro gametocyte. These gametocytes take roughly 8-10 days to reach full maturity.

When ingested by a mosquito feeding on the infected host, the gametocytes leave the erythrocyte shell and mature into gametes. The macro gamete becomes enlarged and spherical. The micro gamete's DNA divides three times to form eight nuclei. Each nucleus pairs with a flagellum to form a microgamete. A macro gamete is fertilized by a microgamete to form a diploid zygote. The zygote then develops into an ookinete. The ookinete traverses the peritrophic membrane and crosses the midgut epithelium, after which it forms an oocyst. The oocyst matures to form multiple haploid sporozoites. The speed of maturation is dependent on temperature, but takes typically one to three weeks. The immature sporozoites break through the oocyst wall into the haemolymph and migrate to the salivary glands, where they complete their differentiation. The mature sporozoites can then be injected into a vertebrate host to complete the lifecycle.

Vectors transmitting *P. vivax* and *P. falciparum* in Sri Lanka[a]

The four species of *Plasmodium* that infect humans have been found to be able to complete the intra-mosquito part of their lifecycle in about 68 mosquito species, all within the genus of *Anopheles* (Diptera: Culicidae). In Sri Lanka, eight anopheline species (*An. aconitus*, *An. annularis*, *An. culicifacies*, *An. nigerrimus* or *An. peditaeniatus*, *An. subpictus*, *An. tessellatus*, *An. vagus*, and *An. varuna*) have been found to be infected with *P. vivax* and nine anopheline species (*An. annularis*, *An. barbirostris*, *An. culicifacies*, *An. pallidus*, *An. peditaeniatus*, *An. subpictus*, *An. tessellatus*, *An. vagus*, and *An. varuna*) have been found to be infected with *P.*

---

[a] Vector ecology is discussed in more detail in Chapter 3 in the subsection "Environmental changes and vector breeding" and in Chapter 4 in the section "Vector ecology"

falciparum. For one infected species (*An. nigerrimus*), the *Plasmodium* species was not specified [2]. The suitability of a mosquito species to act as a vector for *P. vivax* or *P. falciparum* depends not only on mosquito-parasite interactions, but also on mosquito life expectancy (as the *Plasmodium* parasite needs the mosquito to survive the extrinsic incubation period; the period necessary for *Plasmodium* to mature inside the mosquito, which varies with the *Plasmodium* species and depends on ambient temperature), on mosquito feeding frequency, and on mosquito preference for human blood (anthropophily). For a mosquito species to be important in malaria transmission, apart from being a suitable vector, it also has to be (at least seasonally) abundant. Abundance of a vector species is for a large part determined by the ecological suitability of the environmental conditions, which includes the presence and quality of breeding sites, the presence and quality of resting sites, and the access to hosts for taking blood meals. The suitability of a set of environmental conditions varies with the ecological preferences of the vector species, and the environmental conditions are subject to (seasonal and non seasonal) change. In Sri Lanka, *Anopheles culicifacies* is the main mosquito species associated with malaria epidemics [3]. This species has a relatively strong preference to feed on humans (over *e.g.* cattle) compared to the other anopheline species present in Sri Lanka [3], although they are primarily zoophilic (feed mostly on cattle). *Anopheles culicifacies* breeds mainly in pools formed in river and stream beds, and therefore, its density is mostly dependent on temporal and spatial variations in rainfall and river flow. *Anopheles culicifacies* also breeds in abandoned gem mining pits, agricultural wells and to a lesser extent in pools in agricultural water reservoirs [2].

The influence of weather on *Plasmodium* transmission

Weather-related variables such as temperature, humidity, wind, and rainfall form an important part of the set of the environmental conditions that influence mosquito population dynamics and biting behaviour, and hence the transmission of *Plasmodium*. Temperature influences the speed of mosquito (population) development, including the frequency with which blood meals are taken [4]. Humidity is important for the life expectancy of adult mosquitoes, as they are prone to desiccation. Wind may increase desiccation and hamper host seeking. Rainfall influences temperature and relative humidity, but above all provides water necessary for breeding, as the immature stages of the mosquito life cycle are aquatic. The

optimum amount and frequency of rainfall depends on the physical nature of the breeding site and the requirements of the vector species, and rainfall may act indirectly on mosquito breeding when it occurs upstream or when water is transported for irrigation purposes. Hence the impact of rainfall on mosquito population dynamics is complex; A large amount of rainfall at once may flush out breeding sites and wash away adults, while continuous low volume rainfall may not be optimal for colonizing mosquito species that require temporary breeding sites.

Apart from influencing mosquito population dynamics and biting behaviour, temperature also influences the length of the extrinsic incubation period of the *Plasmodium* parasite (The process of sporogony ceases below 16 °C and above 40 °C [5,6]).

Spatial patterns of malaria in Sri Lanka

The spatial variation in annual precipitation has been linked to spatial variation in malaria endemicity in Sri Lanka by early malariologists who used a classification of the country into a wet, intermediate and dry zone [7] based on the amount of rainfall received during the south-west monsoon. The region receiving the most annual precipitation has the least malaria, and endemicity increases with decreasing annual rainfall. The fact that the districts in the extreme south west of the island (Galle and Kalutara) have always been virtually free of malaria is attributed to the wet climate in which rivers flow year round without pooling. In the south west, only a drought might cause pooling in rivers and hence create conditions suitable for the breeding of *An. culicifacies*. For example, districts with wet and intermediate annual rainfall in this region have repeatedly been affected by malaria epidemics, mostly attributed to droughts due to a failing south-west monsoon (which occurs normally between February and July), while districts towards the north and east with dryer climates (and with a higher malaria endemicity) were less affected [7]. In contrast, towards the north and east, where the climate is much dryer (particularly during April – September), rivers often run dry, and rainfall creates new puddles, especially following a period of drought.

The district of Nuwara Eliya, in the hills situated in the south-centre of the country, is also virtually free of malaria, and the few cases recorded there probably resulted from infective bites received elsewhere. With increasing altitude, temperature decreases

and thus at higher altitude conditions are unfavourable for malaria transmission. Other factors influencing spatial distribution of malaria are malaria control interventions (some areas are less accessible, *e.g.* during conflict), population density and variable status of economic development. With increasing wealth, the quality of housing and the level of personal protection against mosquito bites increases.

Temporal patterns in malaria time series

Malaria in Sri Lanka shows a strong seasonality, which varies from bimodal in the south-west, to virtually monomodal in the extreme north and east. Rainfall seasonality follows the same spatial pattern, with a bimodal seasonality in the south-west, and a virtual monomodal seasonality in the north and east.

Apart from annual seasonality, a five-year periodicity was observed by Gill in data on epidemics from 1906 to 1934 [8]. Epidemics in the period from 1867 to 1943 (before large scale vector control was implemented) were found to be significantly associated with El Niño Southern Oscillation (ENSO) [9]. Epidemics were more likely to occur during El Niño years, when the southwest monsoon is was likely to be less intense and involved less rainfall. However, in more recent years, the direction of the association between precipitation and ENSO has been reversed [10]. ENSO could thus still be used to predict rainfall, and possibly malaria, if the link between rainfall and malaria persists.

Immunity

Due to the unstable malaria transmission pattern, it is assumed that the population in Sri Lanka does not have high levels of acquired anti-disease and anti parasite immunity, unlike the situation in highly endemic regions in sub-Saharan Africa [2]. Therefore, both adults and children suffer from the disease, with few asymptomatic carriers. However, some age-related immunity may exist, since adults have a slightly lower incidence than children while they are likely to be more exposed [11,12]. In a study in the area of Kataragama in the south, in adults, higher levels of antibodies to epitopes on circumsporozoite protein were found than in children, and this was associated with a higher frequency of inoculations in adults [13]. However, in the same study, no correlation was found between seroconversion and malaria infection in individuals, leaving the role of these antibodies in protective immunity unclear. Also, in a study in a village in Anuradhapura District, high levels of prevalence (97%) of

antibodies to epitopes on circumsporozoite protein were found during the peak of the transmission season, but these antibody concentrations and prevalence decreased rapidly with time after the end of malaria transmission, and antibody concentrations did not correlate with the presence of blood-stage malaria infections, thus leaving the protective effect unclear [14]. When protective immunity is high as a result of high levels of transmission, as is the case of many African settings, it strongly confounds the relationship between (all age) incidence time series and transmission [15]. When immunity plays an important role in disease transmission, it may create / maintain so called "endogenous" cycles in incidence time series, even when the vectorial capacity (the vector population's potential to transmit malaria) is at a constant level. Cyclical patterns in incidence time series may thus partly be caused by dynamics of immunity, and this may confound the relationship between incidence and extrinsic drivers such as weather [16]. Because of extremely low sporozoite rates in vector mosquitoes in Sri Lanka, it is difficult to measure the entomological inoculation rate (EIR), which is otherwise a good measure of the risk of inoculation. In the absence of reliable estimates of EIR, the malaria case incidence might be a satisfactory measure of parasite transmission, provided that protective and anti disease immunity is low.

Brief history of malaria and malaria control in Sri Lanka

The decline of the irrigation-based civilisation with its centre in Polonnaruwa (in the northern, dry zone of the country) at the end of the twelfth century has been attributed to the introduction of malaria from South India [17], although it is also well possible that the decline of the civilisation was triggered by destruction of the infrastructure by (human) foreign invaders. The Portuguese and Dutch, who successively occupied the coastal areas (1505–1658 and 1658-1798, respectively) mention unhealthiness of certain areas because of periodic fevers [18]. The British, who colonized Sri Lanka from 1798 to independence in 1948, reported more frequently on a malaria-like illness including a report of an epidemic in 1803. In 1861, the British planted cinchona in Sri Lanka. Administration of quinine, derived from cinchona bark, remained the sole antimalarial control activity until 1921, when vector control in the form of environmental management of breeding sites, oiling and larvivorous fish was started [19]. In 1867, the Civil Medical Department started systematic reporting, and annual records of causes of death, including those attributed to malaria, are available since 1911, a year after the establishment of the Anti Malaria Campaign (AMC). The most

serious malaria epidemic recorded in the history of Sri Lanka was the 1934 – 1935 epidemic, which claimed 80 000 deaths. The Malaria Control Programme began in 1945 with DDT indoor residual spraying (IRS), which was associated with a 100-fold reduction in morbidity and mortality over ten years. The Malaria Eradication Programme was started in 1958, and blanket IRS brought malaria down to seventeen cases in 1963. DDT spraying was ceased in 1964 and in 1967/1968, a "post eradication" *P. vivax* malaria epidemic occurred. DDT spraying was then re-introduced, but after the discovery of DDT resistance in 1969 [20], it was replaced by malathion spraying in 1973. Thereafter, malaria morbidity levels fluctuated with epidemics in 1975 and 1986 when *falciparum* malaria morbidity levels were especially high. Mortality rates since 1960 have, however, remained lower than any time since recording started [21]. After decentralization of the AMC in 1989, the spraying program was revised, and only selected malarious areas were sprayed in a blanket approach. After 1998, for IRS, malathion was mostly replaced by other insecticides such as lambda-cyhalothrin, deltamethrin, and fenitrothion. In 1993, the WHO Global Malaria Control Strategy [22] was adopted, which included integrated and selective vector control with targeted spraying only, and distribution of insecticide treated bednets. Also, early diagnosis and prompt treatment was emphasised. The two other main elements of the Global Malaria Control Strategy were early detection, containment or prevention of epidemics, and to strengthen local capacities in research to permit the regular assessment of the malaria situation and its ecological, social and economic determinants. In 1999, the Sri Lanka government approved of the Roll Back Malaria (RBM) initiative. Malaria control activities are currently funded by the Sri Lankan government and the Global Fund to fight Aids, Tuberculosis and Malaria (GFATM).

Thus, in Sri Lanka, organized efforts to reduce transmission of *Plasmodium* species that infect humans have focussed on reducing the availability of the parasite in the human population, by administrating antimalarial drugs to patients after confirmation of the presence of *Plasmodium* parasites, and on vector control[b]. Vector control has focused on reducing the mosquito population and on reducing the mosquito life expectancy. This was done by reducing the presence and suitability of breeding sites

through environmental management [23], larviciding and distributing larvivorous fish, by reducing the quality of resting sites by IRS of insecticides, which affects the life expectancy of mainly indoor resting (endophilic) mosquito species, and by increasing the risk associated with taking a blood meal from human hosts, through distribution of insecticide treated bed nets which form a physical barrier to, repel, and kill indoor biting (endophagic) mosquitoes (Note that IRS also repels and kills endophagic mosquitoes). Because *An. culicifacies* is relatively endophilic, IRS is relatively effective against this species.

Other potential tools for *Plasmodium* transmission control, such as sterile insect technique [24,24], or vaccination, have not been applied in Sri Lanka.

Objectives of this thesis

This thesis was motivated by the need of the Anti Malaria Campaign of Sri Lanka for malaria risk maps and malaria case number predictions to assist in the planning for malaria control. The AMC of Sri Lanka has been successfully combating malaria using integrated and selective vector control and early diagnosis and prompt treatment. However, despite a wealth of high resolution data collected over decades, a malaria forecasting system was not in place, and there was a need for detailed island-wide maps of malaria incidence to permit the assessment of the malaria situation and its determinants. The overall aim of this thesis was to describe the spatial and seasonal distribution of malaria in Sri Lanka and associated factors, and to develop a malaria forecasting system.

The specific objectives were:

1 To describe the spatial variation of malaria in Sri Lanka.

2 To estimate the risk of a malaria epidemic associated with a tsunami natural disaster in Sri Lanka.

3 To evaluate the effect of a tsunami disaster on malaria transmission in Sri Lanka, and to evaluate the response of national and international organisations to the tsunami in terms of malaria relevant actions.

---

[b] Vector control is discussed in more detail in Chapter 3, in the subsection "Vector control strategies and insecticide resistance", and Chapter 4 in the section "Vector control and personal protection since the tsunami"

4 To quantify the correlation in space between seasonality of malaria and seasonality of rainfall in Sri Lanka.

5 To assess the relationships between monthly malaria case count data series and monthly mean rainfall series in Sri Lanka.

6 To develop and implement a model for short term malaria prediction in Sri Lanka.

7 To develop a statistical methodology for analysing over dispersed temporal count data and implement this methodology for short term malaria prediction.

Chapter 2 – Sri Lanka Malaria Maps

This chapter was published in the Malaria Journal 2003, 2:22

Olivier J.T. Briët[1*], Dissanayake M. Gunawardena[2], Wim van der Hoek[1], Felix P. Amerasinghe[1]

[1]International Water Management Institute, Colombo, Sri Lanka

[2]Anti Malaria Campaign, Badulla, Sri Lanka

*corresponding author. E-mail: o.briet@cgiar.org

Abstract

Background

Despite a relatively good national case reporting system in Sri Lanka, detailed maps of malaria distribution have not been publicly available.

Methods

In this study, monthly records over the period 1995 – 2000 of microscopically confirmed malaria parasite positive blood film readings, at sub-district spatial resolution, were used to produce maps of malaria distribution across the island. Also, annual malaria trends at district resolution were displayed for the period 1995 – 2002.

Results

The maps show that *Plasmodium vivax* malaria incidence has a marked variation in distribution over the island. The incidence of *Plasmodium falciparum* malaria follows a similar spatial pattern but is generally much lower than that of *P. vivax*. In the north, malaria shows one seasonal peak in the beginning of the year, whereas towards the south a second peak around June is more pronounced.

Conclusion

This paper provides the first publicly available maps of both *P. vivax* and *P. falciparum* malaria incidence distribution on the island of Sri Lanka at sub-district resolution, which may be useful to health professionals, travellers and travel medicine professionals in their assessment of malaria risk in Sri Lanka. As incidence of malaria changes over time, regular updates of these maps are necessary.

Background

The Anti Malaria Campaign (AMC) Directorate of the Ministry of Health in Sri Lanka maintains a relatively good national case reporting system. However, maps of malaria disease distribution over the island have not been available to a wide public, until a recent publication of a map based on 1989-1994 incidence data at district resolution [2]. Travel medicine Internet sites describe in their advice to travellers to Sri Lanka merely that the risk of malaria is present all year round in all areas (below 800 m altitude), except in the districts of Colombo, Kalutara, and Nuwara Eliya, and sometimes unrealistic maps are posted.

In Sri Lanka, two species of malaria, *Plasmodium vivax* and *Plasmodium falciparum*, are present. The main vector is *Anopheles culicifacies*, which breeds mainly in pools in stagnant rivers, and therefore, its density is mostly dependent on temporal and spatial variations in rainfall and river flow. *Anopheles culicifacies* also breeds in abandoned gem mining pits and agricultural wells. Vectors of less importance are *Anopheles annularis*, *Anopheles subpictus*, *Anopheles tessellatus* and *Anopheles vagus* [2].

This publication provides information on spatial and temporal distribution of malaria incidence on the island of Sri Lanka. Malaria incidence maps are useful in allocating limited malaria control resources to the malaria prone areas at the right time. They may also be useful to health professionals, travellers and travel medicine professionals in their assessment of malaria risk in Sri Lanka.

Methods

The mapping is based on monthly records over the period January 1995 – December 2000 of microscopically confirmed malaria parasite positive blood film readings, at the spatial resolution of Medical Officer of Health (MOH) areas. These were collected by the AMC from aggregated disease records reported by governmental hospitals and mobile clinics. MOH area boundaries are in accordance with the Divisional Secretariat Division (DSD) boundaries (See additional files 1a and 1b: Map and list of Divisional Secretariat Divisions), except that some MOH areas cover multiple DSDs. DSDs are administrative units below the district level with a median population of about 50,000 and an average surface of 208 km^2. District resolution 2001 and 2002 data were included to show recent developments.

Most people in Sri Lanka with suspected malarial fever seek diagnosis and treatment in government health facilities [2]. In all provincial hospitals and in malaria endemic zones also in district and rural hospitals, and in some dispensaries, a microscopist is permanently available for laboratory diagnosis of malaria. Private clinics usually have limited facilities or expertise available for malaria detection, except the private hospitals in Colombo. When parasites are detected, patients are treated with chloroquine 10mg/kg bodyweight, and normally with 8-amino quinaline (primaquine) against liver stages of *P. vivax*.

In the few cases where records for one month or two succeeding months were missing (due to absence of a microscopist), data were estimated by interpolation of monthly case series. In situations where malaria confirmed case data for three or more succeeding months were not available, these months' data were marked as missing.

As a denominator for the incidence calculations, population estimates (See additional file 2: Population) were made by exponential interpolation (and extrapolation to 2002) of 1994 and 2001 census data from the Department of Census and Statistics (http://www.statistics.gov.lk/Documents/census2001/resultindex.htm). For those districts in the north and east not covered by the census, and for which only a district total estimate was posted, DSD populations were estimated according to the population distribution over the districts from data posted by the North East Provincial council (http://www.nepc.lk/index.htm).

The GIS package ArcView was used to modify a DSD map of Sri Lanka to MOH area boundaries and ArcView and MapInfo were used to produce maps of malaria distribution across the island.

Results and discussion

There are several concerns with the quality of the data. In the North and East, malaria case data from there may be grossly underestimated. Due to the armed conflict there was shortage of trained microscopists in these areas and only a small part of the clinical cases is microscopically confirmed [25]. In the rest of the country, the availability of field assistants for blood film collection and the availability of microscopists was high, and the authors estimate the proportion of microscopically confirmed cases to be about 70% [26]. Unfortunately, we have no precise data available to study the effect of the availability of field assistants and microscopists on

the number of blood films examined. In general, there is high acceptance of blood filming by the population [27,28].

It is AMC policy to cross-check 10% of *Plasmodium* positive blood films, and 10% of negative films for parasite presence and species identification, both at District and Central levels. However, after decentralisation of the AMC in 1989, cross-checking was often not performed. Only sporadically blood films were cross-checked at the central laboratory, and no records were kept. In June 2000, a new policy was installed to cross-check films at the central laboratory. Mostly films with doubtful readings were sent to the AMC central laboratory for cross-checking, and only from a limited number of districts and months. Therefore, we could not estimate the error rate for the period under study. An AMC report over the year 1988, before the decentralisation, states a species misidentification of 0%, an error of 0.2% false positives (1.6% of positive slides cross-checked), and 1% false negatives (5% of negative slides cross-checked) [29]. We believe that the quality has since improved as microscopists received more extensive training (1 year versus 6 months) since 1990.

Self-treatment with anti-malarials is relatively uncommon in Sri Lanka. In four-hundred-and-forty-three household interviews in 1992 in Kataragama, Moneragala District, none reported keeping a stock of anti-malarial drugs at home (DMG, unpublished data). In a survey in 1999 at governmental hospital level in nine malarious districts (outside the conflict area), none out of nine-hundred patients diagnosed with *P. falciparum* reported the use of anti-malarial drugs prior to presentation at the hospital, whereas 19% had taken non anti-malarial drugs, mostly administered by the government hospital or dispensary [30]. However, in 2000 in Mallavi, Mullaitivu District (in the conflict area), 7.4% of patients reported self-treatment with chloroquine prior to presentation to the outpatient department (OPD), and 84.5% with non anti-malarial drugs [25]. It is not known how many people successfully treated themselves with anti-malarial drugs and therefore did not present themselves to the governmental facility in the latter two studies.

Patients who seek treatment at non-governmental health facilities are not registered, and this leads to further underestimation of the number of cases. In a study in three MOH areas in Moneragala, only about half the cases were treated at governmental health facilities and therefore registered, with considerable variation at Grama

Niladhari resolution [26]. Grama Niladhari are administrative units with the highest spatial resolution used in Sri Lanka. However, at coarser resolution, gross spatial bias due to treatment at private facilities is expected to be limited, as governmental facilities are the preferred diagnosis and treatment centres (69% in an irrigation resettlement area (Mahaweli System C) in Badulla District[31], >75% in Kataragama, Moneragala District [11], 84% in a location in Hurulawewa, Anuradhapura District [27], 83 – 97% in four villages in and around Lunugamwehera irrigation project, Hambantota District [32], even in the conflict areas (80% in Mallavi, Mullaitivu District [25]).

Another spatial bias is the fact that cases detected in occasional mass blood surveys in selected villages in high risk areas are also included in the statistics. However, these blood films tend not to exceed 1% of the total examined.

Aggregated case records from the health facilities were not corrected for recrudescence of *P. falciparum* or relapse of *P. vivax*. It is, therefore, possible that patients with treatment failure due to incomplete drug compliance or resistance, were recorded more than once, thereby overestimating the incidence. Interviews in Kataragama (Moneragala District) of malaria patients revealed that drug non-adherence is very low (none for forty-three recrudescent cases [33], three for more than seven-hundred-and-twelve cases (<0.4%) [34]. In Malavi (Mullaitivu District), however, interviews revealed 26.2% non-adherence to full treatment, mostly (58%) for reasons of side effects [25]. It is not known in how many of these cases this resulted in treatment failure. In Sri Lanka, no studies have yet employed molecular methods to differentiate between recrudescence and re-infection [35]. Instead, studies at several locations have used different arbitrary time periods between successive infections for classifying a successive infection as recrudescent or as new. Also, some studies have used active follow up methods instead of passive methods to detect recurrent infections. Corrections based on these active detection studies tend to overestimate the number of double counted *P. falciparum* cases, as chloroquine resistant recrudescent infections show less severe clinical symptoms, and therefore have a lower probability of being recorded in the AMC registers [34]. In Colombo, during 1992 – 1993, a study using active detection reported 55% (n = 129) recrudescent cases within 40 days of follow up. However, 61% of these were non patent and these people reported that they would not have sought treatment.

Therefore, only 22% of cases would have been double recorded if detected passively [34]. Handunnetti and colleagues found in a passive case detection study in 1992 in Kataragama (Moneragala District), that 26% (n = 616) of *P. falciparum* episodes occurred within 31 days of the previous episode in the same person [34]. A more recent study using passive case detection during 1998-1999 in Kataragama and Buttala (Moneragala District) found 12% (n = 359) of cases re-occurring within 28 days [33]. An active case-detection follow-up study in 1999 in nine malarious districts (outside the conflict area) found parasites in 34 – 62% (Table 1) of patients within 28 days after diagnosis and treatment of *P. falciparum* [30]. It is interesting to note that there is a strong positive correlation (binomial regression, r = 0.81, p < 0.01, n = 9) between the proportion of recrudescence infections and *P. falciparum* incidence (even if the incidence is corrected by assuming that each recrudescent case is counted twice). Based on this regression one could consider adjusting reported cases of *P. falciparum* in each MOH area, which would bring down higher incidence rates relatively more than lower incidence rates. We did not do so because of likely overestimation of the number of double counted cases by the active detection method used in the follow up study. With regard to *P. vivax* relapses, Fonseka and Mendis [36] estimated a rate of 18% from patients in Colombo during the period 1981 – 1984. These people had acquired their infections elsewhere in the country, and most of them suffered from the relapse within 24 weeks after the primary attack.

Table 1 - Incidence and recrudescence of *Plasmodium falciparum* in 1999 in nine districts of Sri Lanka

District	*P. falciparum* cases	Population	*P. falciparum* incidence (x 1000)	Recrudes-cent cases*	Number of patients followed for 28 days*	Proportion re-crudescent*
Anuradhapura	5,132	725,557	7.07	49	99	0.49
Badulla	633	1,073,134	0.59	34	100	0.34
Hambantota	1,018	495,702	2.05	35	100	0.35
Kurunegala	2,073	1,462,149	1.42	42	100	0.42
Matale	1,116	495,511	2.25	36	100	0.36
Moneragala	7,215	448,226	16.10	62	100	0.62
Polonnaruwa	978	319,632	3.06	50	100	0.50
Puttalam	3,375	843,410	4.00	53	97	0.55
Ratnapura	2,685	1,035,690	2.59	42	100	0.42

* Data reproduced with permission of Dr. G.N.L. Galappaththy [30]

Another concern for data quality is that the population census data may be less reliable in the North and East. Also, an important number of malaria infections may not have been contracted at the place of reporting. This may especially be true for infections contracted by military personnel in the conflict zone and reported in their place of residence while on medical leave. Furthermore, until 2000, cases were generally ascribed to the MOH area of the reporting hospital, regardless of the place of residence of the patient. The year 1998 was the most complete in terms of malaria case records.
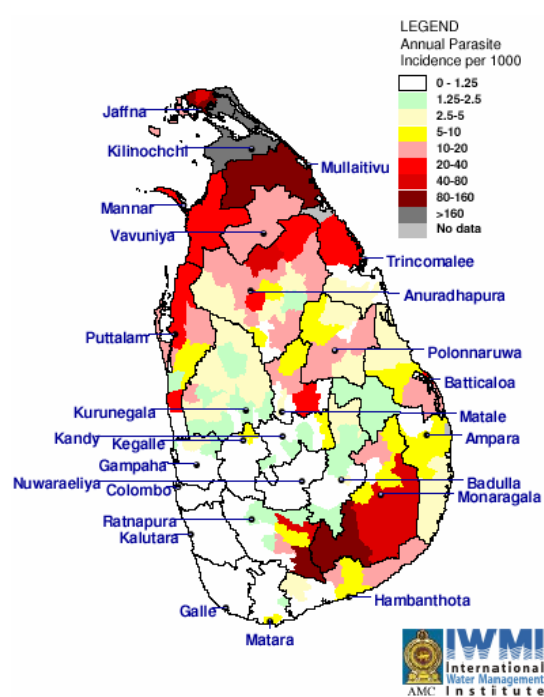
Figure 1 - Annual parasite incidence of *Plasmodium vivax*

Map of the districts of Sri Lanka with annual parasite incidence (API) of *P. vivax* malaria cases at Medical Officer of Health (MOH) area resolution over the year 1998.
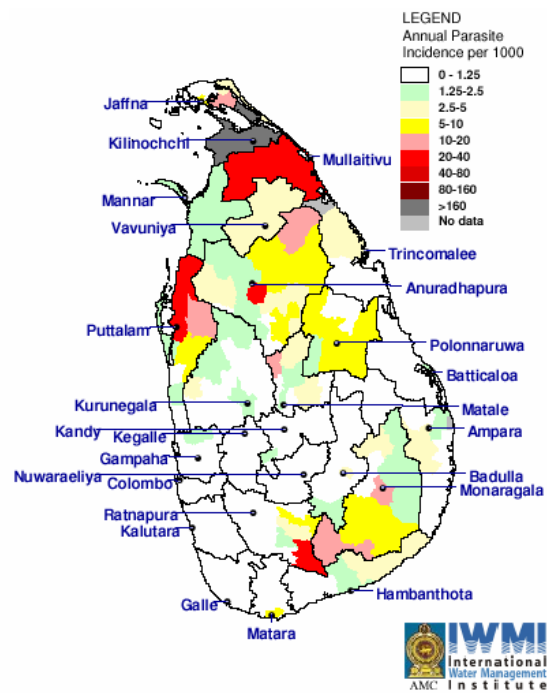
Figure 2 - Annual parasite incidence of *Plasmodium falciparum*

Map of the districts of Sri Lanka with annual parasite incidence (API) of *P. falciparum* malaria and mixed infections of both *P. vivax* and *P. falciparum* at Medical Officer of Health (MOH) area resolution over the year 1998.

Figure 1 shows that the annual parasite incidence (API) of *P. vivax* malaria cases at MOH area resolution had marked variation over the island. Particularly, the districts of Jaffna, Kilinochchi and Mullaitivu in the north, and the district of Moneragala and the south-eastern MOH areas in Ratnapura district show high malaria incidence. The API of *P. falciparum* (and mixed) infections (Figure 2) was generally much lower

than the API of *P. vivax*, although the spatial distribution is somewhat similar. In the districts of Batticaloa and Ampara in the east, the proportion of *P. falciparum* was much lower than elsewhere in the country.

Clearly, the northern areas are facing a serious malaria problem. Difficulties in obtaining prompt treatment may have enhanced malaria transmission. In the rest of the country this factor seems of a lesser importance, as the health systems are generally well developed. Socio-economic factors such as personal protection against mosquitoes and quality of housing construction are important in explaining the distribution of malaria incidence. More important, however, are factors influencing malaria mosquitoes, such as temperature (altitude), rainfall and resulting river flow (See additional files 3: temperature, 4: altitude, 5: rainfall, and 6: rivers and lakes), but also (chemical) control efforts by the AMC. Especially the latter factor has historically played an important role in the malaria epidemiology in Sri Lanka [21]. After 1983, no more governmental vector control has been implemented in the northern areas. To learn more about the relative importance of socio-economic and environmental risk factors for malaria, a spatial regression analysis linking incidence directly to covariates (as information on vector density and distribution is scarce) is being done, the results of which will be disseminated in due course.
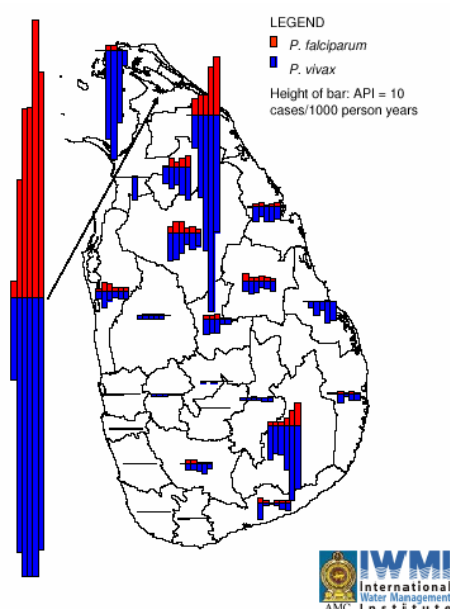
Figure 3 - Trends of annual parasite incidence

Trends of annual parasite incidence of *P. falciparum* (red bars) and *P. vivax* (blue bars) malaria over the years 1995 (bar on far left) to 2002 (bar on far right), at district resolution. The height of the bars in the legend represents an annual parasite incidence of 10 cases per 1000 persons.

Figure 3 shows the trends of annual parasite incidence of *P. falciparum* and *P. vivax* malaria over the years 1995 to 2002, at district resolution. *Plasmodium falciparum* and *P. vivax* generally show similar trends over the 8-year period. However, there is considerable variation in temporal trends over the country, even at a relatively short distance: In Jaffna, malaria incidence went down after 1998, whereas it still increased in the neighbouring districts of Kilinochchi and Mullaitivu. For Mannar, data were incomplete. Although the malaria incidence showed a general increase over the 1995 – 2000 period, it declined strongly after 2000. The ongoing peace process may be an important contributing factor for the recent decline in cases in the conflict zone. Notably the access to the area for spray teams has been increased. Also, foreign aid organisations have been providing insecticide-impregnated bednets in the affected areas. The variation in temporal trends and socio-political developments illustrate the need for regular updates of malaria distribution maps such as shown in Figure 1 and Figure 2.
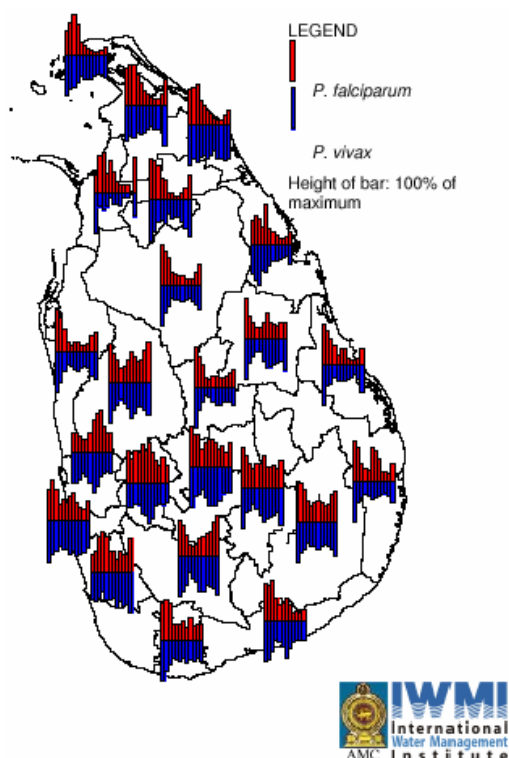
Figure 4 - Geometric mean monthly parasite incidence patterns

Geometric mean monthly parasite incidence patterns of *P. falciparum* (red bars) and *P. vivax* (blue bars) malaria from January (bar on far left) to December (bar on far right), relative to the month with the highest geometric mean incidence, over the period January 1995 to December 2000, at district resolution. The height of the bars in the legend represents 100 percent (The month with the highest geometric mean incidence for the respective malaria species).

Figure 4 shows the geometric mean monthly parasite incidence patterns of *P. falciparum* and *P. vivax* malaria over the period January 1995 to December 2000, at district resolution. In the northern districts, *P. falciparum* peaks around March, whereas towards the south the peak is shifted towards January. *Plasmodium vivax* generally peaks in January. Except for the northern districts, a second malaria peak of either species occurs around July. This peak is especially pronounced in the west-central districts with very low malaria incidence, where it can even outweigh the January peak. Roughly, in all districts, *P. falciparum* and *P. vivax* follow the same seasonal pattern, although *P. vivax'* seasonality is less explicit. This observation could be due to relapses (activation of hypnozoites in the liver). Therefore, the *P. falciparum* incidence seems a better proxy for *P. falciparum* malaria transmission seasonality than *P. vivax* incidence is for *P. vivax* transmission seasonality, although these patterns are also somewhat smoothed by recrudescence.

There is evidence of considerable spatial variation in the risk of malaria transmission at a higher resolution than the MOH area scale presented in this paper. Malaria is a disease of rural areas and cities are mostly unaffected. The distance of houses to breeding sites of malaria vectors within a MOH area is an important risk factor [37,38]. The authors of this study are currently working on a malaria risk map of the Badulla and Moneragala districts at Grama Niladhari resolution.
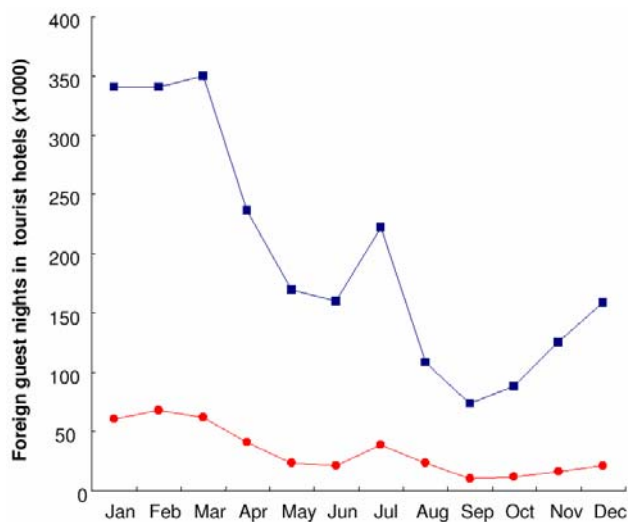
Figure 5 - Foreign guest nights in tourist hotels

Monthly foreign guest nights spent in tourist hotels in 2001 in malarious areas with an annual parasite incidence > 1 case/1000 population (red lines and dots) and non malarious areas (blue lines and squares). Source: Ceylon Tourist Board: Annual Statistical Report 2001 [39]. Note that this figure is

not representative for other years because of the attack on the Bandaranaike International Airport by the Liberation Tigers of Tamil Eelam on 24 July 2001, after which tourism plummeted.

On average, 370,000 tourists visit Sri Lanka annually, of whom the majority (63%) is of European origin (http://www.lanka.net/centralbank/y-td_tourism.html). Roughly 14% of tourist hotel nights booked by foreigners in 2001 [39] was in areas with a risk of malaria (API > 1 case/1000 population). Of these, most were spent during months of transmission, as the tourist seasons coincide with the inter-monsoon periods (Figure 5) when malaria transmission is at its maximum. Some of the important tourist destinations, such as the ancient cities of Anuradhapura, Polonnaruwa, and Sigiriya, and the Yala and Uda Walawe national parks are situated in endemic areas, but these are mainly popular for day trips. Most tourists will therefore not be exposed during evening or night time, when *An. culicifacies* is most active. Tourist hotels generally provide anti mosquito measures such as pyrethrum mosquito coils and bed nets, and most hotel rooms have a fan or air conditioning, so contact with nocturnal indoor-biting vectors is limited. Repellents are recommended when outdoors after dusk. There is no justification for prescribing chemoprophylaxis to tourists who intend to remain in resorts in the non-malarious areas and make only day trips to destinations in the malaria endemic areas. Physicians and travel clinics should tailor their advice on prophylactic drugs to the individual traveller, taking into account the itinerary and time of travel. The AMC advises travellers to malaria endemic areas (with an API of *P. falciparum* and/or *P. vivax* above 10 per 1000 population) to take a weekly dose of 300 mg chloroquine (for adults) as prophylactic measure from two weeks before the visit until four weeks after the visit. In case of treatment failure due to chloroquine resistance, sulfadoxine / pyremethamine is available at all governmental health facilities in the endemic areas. Carrying anti-malarial drugs for self administration (standby treatment) should not be recommended for Sri Lanka, as facilities for diagnosis and treatment are available in all parts of the country.

Conclusions

This paper provides the first publicly available maps of both *P. vivax* and *P. falciparum* malaria incidence distribution on the island of Sri Lanka at sub-district resolution. The maps show that both *P. vivax* and *P. falciparum* malaria incidence have a marked variation in distribution over the island, even within districts. The incidence of *P. falciparum* malaria follows a similar spatial pattern to that of *P. vivax*

but is generally much lower. In the north, malaria shows one seasonal peak in the beginning of the year, whereas towards the south a second peak around June becomes more pronounced.

These maps may be useful for the planning of malaria control activities. They also may be useful to health professionals, travellers and travel medicine professionals in their assessment of malaria risk in Sri Lanka. However, as incidence of malaria changes over time, regular updates of these maps are necessary.

List of additional files to this chapter (printed in Appendix)

Additional file 1a – Divisional Secretariat Divisions

Map of Divisional Secretariat Divisions in Sri Lanka. Source: The Survey Department, Sri Lanka.

Additional file 1b – Divisional Secretariat Divisions

DSDs.xls: List of Divisional Secretariat Division names and their number corresponding to the number printed in additional file 1a.

Additional file 2 – Population

Map of population by divisional secretariat division in Sri Lanka. One dot represents 1000 people. Sources: Department of Census and Statistics

(http://www.statistics.gov.lk/) and North East Provincial Council
(http://www.nepc.lk/index.htm).

Additional file 3 – Temperature

Maps of monthly average temperature in Sri Lanka. Source: IWMI World Water and
Climate Atlas (http://www.iwmi.cgiar.org/WAtlas/atlas.htm).

Additional file 4 – Altitude

Map of altitude in Sri Lanka. Colour shading as recommended by the GLOBE project
(except white background). Green contour line: 500m, black contour lines: 1000m.
Source: GLOBE project (http://www.ngdc.noaa.gov/seg/topo/globe.shtml).

Additional file 5 – Rainfall

Maps of monthly rainfall (75% probability) in Sri Lanka. Source: IWMI World
Water and Climate Atlas (http://www.iwmi.cgiar.org/WAtlas/atlas.htm).

Additional file 6 – Rivers and lakes

Map of river network and lakes / tanks in Sri Lanka. Source: The Survey Department,
Sri Lanka.

Addendum to Chapter 2 – Regression analysis of spatial distribution of malaria

Part of these result were presented at the 58th Annual Session, Sri Lanka Association of the Advancement of Science [40]

Background

In chapter 2, the spatial distribution of malaria in 1998 at sub-district level was presented. Although some covariates influencing malaria distribution were discussed, this discussion was not supported by a statistical analysis. This chapter provides the spatial statistical analysis of malaria incidence over 1998 in relation to some covariates

Materials and methods

Malaria data were the aggregated clinical malaria case numbers over 1998 at Medical Officer of Health (MOH) area spatial resolution, due to both *P. vivax* (Figure 1) and *P. falciparum* (Figure 2) infection. In 1998, there were 226 MOH areas in Sri Lanka. Two of these were excluded because of missing malaria data in 1998. For each MOH area, the mid year population for 1998 was calculated by (temporal) interpolation of census data, which were collected by the Department of Census and Statistics [41] and North East Provincial Council [42].

Using a geographic information system (GIS), a digital map of Divisional Secretariat Division (DSD) boundaries (See additional files 1a and 1b: Map and list of Divisional Secretariat Divisions), was modified into a map of MOH areas by aggregating some DSD areas. The digital MOH area map was then used to calculate the surface area for each MOH area, and to extract data from GIS data maps. Thus, data for the following variables were obtained: the mean altitude [43], mean rainfall [44], a binary indicator for government "cleared" (controlled) areas (as opposed to areas that are not under government control because of conflict) if more than 33% of the surface of an area was "cleared" [45], proportion of area surface covered with sparse forest [46], proportion of area surface covered with dense forest [46], proportion of area surface covered with paddy [46], and proportion of area surface covered with a buffer zone of

100 meter around river and tank (reservoirs) shore [46]. The population density was calculated by dividing the population by the surface of the area.

First, each covariate, untransformed and logarithmically transformed (if possible), was tested individually in a non spatial negative binomial model. The residual spatial correlation was calculated using the Moran's I index for spatial correlation on the Pearson residuals. The covariates for which the posterior distribution of the coefficient did not include zero in the non spatial negative binomial were all entered into a multiple Poisson conditional autoregressive ($CAR(\gamma, \tau_s)$) model. In this model, unexplained extra-Poisson variation (variation that is larger than the conditional mean) is assumed to be either due to effects of unobserved covariates, or spatial autocorrelation due to the contagious nature of the disease. The extra-Poisson variation is then modelled through random effects, which can be spatially auto correlated. The modelling of the (spatial) random effects typically enlarges the support of the posterior distribution of the coefficients [47].

Another way of modeling count data is through the negative binomial distribution. The specification for the negative binomial distribution of an independent malaria count y in area i, is $y_i \sim NegBin(\lambda_i, r)$ with the mass function

$$f(y_i \mid \lambda_i, r) = \frac{\lambda_i^{y_i}}{y_i!} \frac{\Gamma(r + y_i)}{\Gamma(r)(r + \lambda_i)^{y_i}} \frac{1}{\left(1 + \frac{\lambda_i}{r}\right)^r}, \text{ which has a conditional mean } \lambda_i \text{ and}$$

variance $\lambda_i + \dfrac{\lambda_i^2}{r}$. As $r \to \infty$, the limiting form of the negative binomial distribution

is the Poisson distribution with mass function $g(y_i \mid \lambda_i) = \dfrac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!}$. The parameter r

thus models extra-Poisson dispersion.

The parameter λ_i was modelled with a logarithmic link function, $\lambda_i = \exp(\mu_i)$ with the logarithmically transformed population as offset.

We considered $\mu_i = \mathbf{x}_i^T \boldsymbol{\beta} + \varphi_i$ where $\boldsymbol{\beta}^T = (\beta_0, \beta_1, \beta_2, ..., \beta_v)$ is a vector of coefficients for $\mathbf{x}_i^T = (x_0, x_{1,i}, x_{2,i}, ..., x_{v,i})$, a vector of an intercept multiplier (usually taken as $x_0 = 1$) and v covariates, and φ_i is an area-specific random effect taking into account the spatial correlation introduced by the spatial structure of unobserved covariates.

For the φ_i's, a CAR(γ, τ_s) model was adopted which assumes that

$$\varphi_i \mid \varphi_{-i}, \tau_s, \gamma \sim N\left(\frac{\gamma \sum \varphi_{-i}}{n_i}, \frac{1}{n_i \tau_s}\right)$$ where γ is a spatial correlation parameter, n_i is the

number of areas bordering area i and $\sigma_s^2 = \dfrac{1}{\tau_s}$ measures between area variation. To

facilitate model fit a Bayesian modelling framework was used. The following prior

distributions were chosen for the parameters: $\beta_0, \ldots, \beta_v \sim U(-\infty, \infty)$, $\tau, \tau_s \sim Ga(0.005, 0.005)$, $\gamma \sim U(a, b)$ with limits a, b specified as described by Gelfand and Vounatsou

[48]. For the (non spatial) negative binomial models, the dispersion parameter of the

negative binomial distribution was assigned a $\psi \sim Ga(0.001, 0.001)$ prior

distribution. The CAR(γ, τ_s) was considered as a prior distribution for the φ_i's. The

effect of this prior distribution is to shrink the observed value of an area to that of the

local mean, where the local mean is the mean of all contiguous areas excluding the

area itself. The posterior distribution of the value of an area is therefore a

compromise between the prior, which is based on the value of neighbouring areas,

and the data for the area. Bayesian CAR models have been widely used in malaria

mapping [49-52]. A non spatial model was also applied. The deviance information

criterion (DIC) [53] was used to determine the best fitting model. The models were

estimated using a Markov Chain Monte Carlo process using three chains. The number

of iterations and the length of the burn-in varied on the models, depending on the

convergence (assessed by studying plots of the Gelman-Rubin convergence statistic as

modified by Brooks and Gelman [54]) and the amount of autocorrelation in the

individual chains. The models were implemented with the software package

"WinBUGS" [55], called from the statistical environment "R" [56] using the package

"R2WinBUGS".

Table 2 - Variables tested as covariate in negative binomial non spatial models.

Explanatory variable	Coefficient			Moran's I			DIC
	Mean	95% CI Lower bound.	95% CI Upper bound.	Mean	95% CI Lower bound.	95% CI Upper bound.	
None				0.61	0.60	0.61	2778
Annual rainfall (m)	-4.02	-4.78	-3.29	0.42	0.38	0.48	2698
Population density (persons / m^2), logarithmically transformed	-0.83	-1.10	-0.57	0.69	0.67	0.69	2749
Altitude (m)	-3.12	-4.09	-1.95	0.59	0.58	0.60	2757
Cleared areas (m^2 / m^2)	-2.05	-3.39	-1.01	0.53	0.45	0.59	2763
River and tank shore (m^2 / m^2), logarithmically transformed	-0.46	-1.03	-0.16	0.36	0.22	0.49	2767
Paddy cultivation (m^2 / m^2), logarithmically transformed	0.22	0.01	0.39	0.61	0.60	0.61	2775
Dense forest (m^2 / m^2)	1.76	0.28	3.66	0.63	0.61	0.65	2775
Sparse forest (m^2 / m^2)	3.91	-1.14	9.99	0.62	0.60	0.62	2778

Legend: Transf. = Transformation, Dev. = Deviance, NA = Not Available, Log. = logarithmically, CI = Credible interval

Results

The results of the (non spatial) negative binomial model, obtained using three Markov chains of 2 000 iterations, including a burn-in of 1 000 iterations (Table 2), showed that the posterior distribution of the coefficient did not include zero for most of the covariates. The mean of the posterior distribution of the dispersal parameter r of the negative binomial was below one, indicating that the malaria cases, conditional on the population and each of the tested covariates, showed a strong extra-Poisson (over) dispersion. The model with the lowest DIC was the model with the total annual rainfall as covariate. Although this covariate lowered the Moran's I, as compared to that of the raw data, it was still significant. If a model shows spatial autocorrelation in the residuals, the posterior density distributions of the coefficients are typically too narrow, and the estimates of the mean of the coefficients may be biased.

In the Poisson models with CAR, the Markov chains of the samples for the coefficient of the covariates showed strong autocorrelation. For reliable estimation, many iterations were thus required. The multivariate model was estimated using three Markov chains of 2 600 000 iterations, including a burn-in of 100 000 iterations. The posterior distributions of the coefficients for annual rainfall, logarithmically

transformed population density and altitude did not include zero, whereas they did for cleared area, logarithmically transformed river and tank shore, logarithmically transformed area of paddy cultivation, dense forest or sparse forest (Table 3).

It should be noted that the deviance and the DIC of the Poisson CAR model was much lower than the deviance and DIC of the best non spatial negative binomial model.

Table 3 - Regression results of malaria incidence in the spatial Poisson CAR(γ, τ_s) model.

Parameter	mean	95% CI Lower bound.	95% CI Upper bound.
Intercept	-9.49	-13.58	-5.63
$\beta_{annual\ rainfall}$	-2.41	-3.28	-1.47
$\beta_{log(population\ density)}$	-1.06	-1.52	-0.65
$\beta_{altitude}$	-2.92	-4.67	-1.31
$\beta_{cleared\ areas}$	0.83	-0.90	2.60
$\beta_{log(shoreline)}$	0.17	-0.09	0.51
$\beta_{log(paddy)}$	0.11	-0.05	0.26
$\beta_{dense\ forest}$	-0.94	-3.42	1.67
γ	0.81	0.54	0.97
τ	0.04	0.04	0.05
Moran's I	0.02	-0.07	0.11
deviance	1419	1380	1465
DIC	1610		

Discussion

The most important variable in explaining the spatial variability of malaria in single variable negative binomial analysis was annual rainfall, with a negative coefficient. As discussed in chapter 6, the spatial distribution of annual rainfall has since long been associated with the spatial distribution of malaria endemicity. The explanation for this is that *Anopheles culicifacies*, the main malaria vector in Sri Lanka, breeds by preference in puddles in stagnating rivers. In the region with the highest annual rainfall, with a strong bimodal rainfall pattern, rivers rarely stagnate, whereas in the region with less annual rainfall (and only one rainfall peak), rivers stagnate during the long dry season.

The second most important variable was (logarithmically transformed) population density, also with a negative coefficient. Population density is an interesting variable with many aspects. It is possible that people have migrated to, or thrived better in areas where environmental conditions were less favourable for malaria transmission, because of malaria [57]. The spatial distribution of malaria would then be an

explanatory variable for the distribution of population density, rather than the inverse. On the other hand, in Sri Lanka, population density is associated with wealth, the wealthier living in cities. The wealthier tend to be better protected from mosquito bites due to improved housing quality, use of electrical fans or air conditioning and purchasing power of bed nets and mosquito repellents. Densely populated places are also associated with environmental pollution, which could affect anopheline mosquito breeding negatively. It is also well possible that the correlation between population density and malaria is completely co-incidental. People have thrived in coastal areas, with income generating activities such as trade, fishing and tourism. The cultivation of many crops is more successful in the region with ample annual rainfall, than in the dryer areas, resulting in a higher population density in these areas. Furthermore, in the hill country, where lower temperatures are not conducive to malaria transmission, there is a large population of tea-estate workers.

The third most important variable was altitude, again with a negative coefficient. It is well known that temperature decreases linearly with altitude, and parasite development in the vector mosquitoes is slow at low temperatures.

The three most important variables in single variable negative binomial analysis (annual rainfall, logarithmically transformed population density and altitude) were all significantly different from zero in spatial Poisson analysis allowing for random effects.

The fact that the variable "safety cleared areas" were not significant in a multivariate analysis could be explained by high collinearity with rainfall and population density.

Chapter 3 – Maps of the Sri Lanka malaria situation preceding the tsunami and key aspects to be considered in the emergency phase and beyond

This chapter was published in the Malaria Journal 2005, 4:8.

Olivier J.T. Briët[1*], Gawrie N.L. Galappaththy[2], Flemming Konradsen[1, 3], Priyanie H. Amerasinghe[1], Felix P. Amerasinghe[1]

[1]International Water Management Institute, P.O. Box 2075, Colombo, Sri Lanka. Tel. +94 11 2787404; Fax. +94 11 2786854

[2]Anti Malaria Campaign Head Office, Colombo, Sri Lanka

[3] Department of International Health, University of Copenhagen, Denmark

*Corresponding author.

Abstract

Background

Following the tsunami, a detailed overview of the area specific transmission levels is essential in assessing the risk of malaria in Sri Lanka. Recent information on vector insecticide resistance, parasite drug resistance, and insights into the national policy for malaria diagnosis and treatment are important in assisting national and international agencies in their control efforts.

Methods

Monthly records over the period January 1995 – October 2004 of confirmed malaria cases were used to perform an analysis of malaria distribution at district spatial resolution. Also, a focused review of published reports and routinely collected information was performed.

Results

The incidence of malaria was only 1 case per thousand population in the 10 months leading up to the disaster, in the districts with the highest transmission.

Conclusion

Although relocated people may be more exposed to mosquito bites, and their capacity to handle diseases affected, the environmental changes caused by the tsunami are unlikely to enhance breeding of the principal vector, and, given the present low parasite reservoir, the likelihood of a malaria outbreak is low. However, close monitoring of the situation is necessary, especially as December – February is normally the peak transmission season. Despite some losses, the Sri Lanka public health system is capable of dealing with the possible threat of a malaria outbreak after the tsunami. The influx of foreign medical assistance, drugs and insecticides may interfere with malaria surveillance, and the long term malaria control strategy of Sri Lanka if not in accordance with government policy.

Background

After the tsunami hit Sri Lanka on 26 December 2004, news reports and public health agencies warned against the possibilities of an increase of vector borne diseases, in particular malaria and dengue. Immediately after the disaster, an estimated 860,000 people were displaced and more than 820 emergency camps established throughout the affected areas [58]. By 14 January, approximately 440,000 people were still sheltered in approximately 460 emergency camps [59]. For maps of the tsunami affected area, we refer elsewhere [60].

Malaria in Sri Lanka is of a highly unstable nature and has historically fluctuated greatly over the years and with significant seasonal differences. Sixty-five to eighty percent of the malaria cases are caused by *Plasmodium vivax* and the remainder by *Plasmodium falciparum* [2]. Recently, an overview of the spatial and temporal distribution of malaria in Sri Lanka over the period 1995 – 2002 was published in this journal [61]. The present publication aims at providing an update on the recent malaria situation, to October 2004 inclusive, and to discuss factors of relevance which may help in assessing the potential of the tsunami and ensuing events to exacerbate the malaria situation.
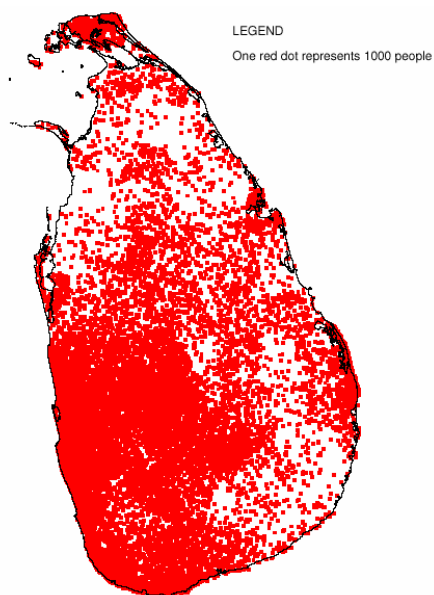
Figure 6 - Population

Map of population by divisional secretariat division in Sri Lanka estimated for mid December 2004. One dot represents 1000 people. Sources: Department of Census and Statistics (http://www.statistics.gov.lk/), North East Provincial Council (http://www.nepc.lk/ and UNHCR (http://www.unhcr.lk).

Methods

Malaria maps were based on monthly records over the period January 2004 – October 2004 (the most recent month for which data recording was complete at the time of writing) of microscopically confirmed malaria parasite positive blood smear readings, at district spatial resolution. These were collected by the Anti Malaria Campaign (AMC) Directorate of the Ministry of Health from aggregated disease records reported by governmental hospitals and mobile clinics. Additionally, in the temporal analysis, monthly data by district for the period 2001 – 2002, and data by sub district for 1995 – 2000 as described by Briët *et al.* [61] were used. The quality of routinely collected information on malaria is described elsewhere [61]. As denominator for the incidence calculations, population estimates for 2001 and beyond were made by exponential interpolation (and extrapolation to December 2004) (Figure 6) as follows. For the districts Mannar, Vavuniya, Trincomalee and Batticaloa, that were not or incompletely enumerated in the 2001 census because of limited access of the government to these conflict affected areas, the 2001 mid-year population was taken from data posted by the North East Provincial Council [42]. For all other districts, the 2001 mid-year population was taken from data posted by the Department of Census and Statistics [41]. The natural annual (mid-2001 to mid-2002 and mid-2002 to mid-2003) population growth rates for Jaffna, Kilinochchi, Mullaitivu, Mannar, Vavuniya, Trincomalee and Batticaloa were taken as the average annual growth rates of all the other districts, calculated from mid year population statistics estimated by the Department of Census and Statistics. For all other districts, these growth rates were calculated for each district separately. For mid 2003 – mid 2004 and beyond, the growth rates for mid-2002 to mid-2003 were used. Further, the number of internally displaced persons (IDPs) was taken into account [62]. For each month and for each district, the net number of immigrants was calculated as the total number of IDPs moved to or within a district since 2001, minus the number of IDPs moved from or within that district. This net number of immigrants was then distributed over the months proportionately to the monthly statistics of IDPs moved to or within a district. Additionally, the number of monthly immigrants from India was taken into account.

A focused review of literature has been performed, identifying crucial information for the outbreak preparedness and control during the emergency phase. The intent was not to present a complete review of malaria in Sri Lanka but to provide information

useful for an assessment of the current situation. A general review of malaria in Sri Lanka can be found in Konradsen, Amerasinghe *et al.* [2].
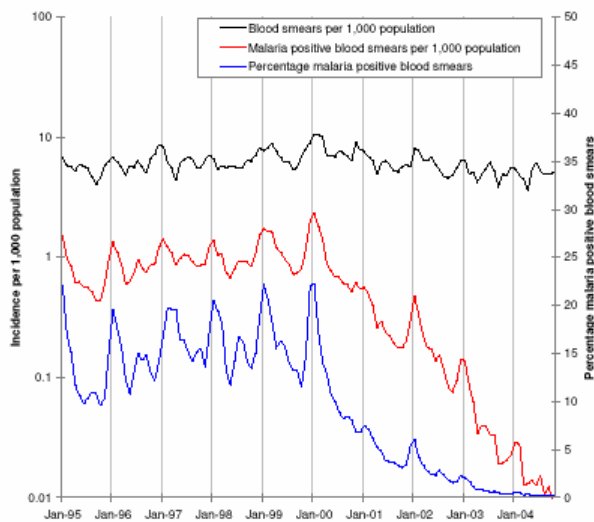
Figure 7 - Monthly parasite and blood smear examination incidence patterns

Monthly parasite incidence patterns of *P. falciparum* and *P. vivax* malaria combined per 1000 population (red line on logarithmic scale), blood smears examined per 1000 population (black line on logarithmic scale), and percentage of blood smears positive for malaria (blue line) from January 1995 to October 2004 in Sri Lanka.
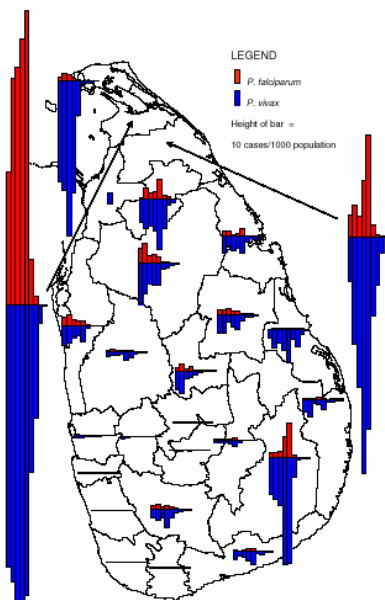
Figure 8 - Trends of parasite incidence

Trends of parasite incidence of *P. falciparum* (red bars) and *P. vivax* (blue bars) malaria over the years November 1995 – October 1996 (bar on far left) to November 2003 – October 2004 (bar on far right), at district resolution. The height of the bars in the legend represents an annual parasite incidence of 10 cases per 1000 persons.

Results and discussion

Present malaria situation and parasite reservoir

The country-wide malaria incidence increased from January 1996 to January 2000, with the typical seasonality of high peaks around January and lower peaks around June – July, but it has decreased dramatically since January 2000 (Figure 7). Figure 8 shows that the recent decrease in the overall malaria incidence in the country is predominantly due to a decrease in incidence in the districts of Vavuniya and Kilinochchi in the north. The decrease was least in the district of Ampara, making it the most malarious district during January – October 2004 (Figure 9 and Figure 10). Although districts on the east coast which were badly affected by the tsunami had been relatively malarious in 2004 as compared to the rest of the country, the maximum of around 1 case per 1000 people over a 10 month period in these districts is remarkably low. The total number of malaria cases in 2003 was 10,510, the lowest since the resurgence of malaria in 1968 when the eradication campaign failed [21]. The year 2004 promises to be three times lower with only 3,037 cases recorded up to October, as opposed to 9,682 cases recorded during January – October 2003. The low incidence is not related to a decline in collection effort, which has decreased only marginally (Figure 7). At the time of writing, malaria incidence information for the months of November and December was still incomplete. In November 2004, without the figures for the non endemic districts Gampaha and Kalutara, and data from a few medical institutions in Mannar and Mullaitivu missing, thus far only 230 cases were recorded. In the malaria endemic districts, December, January and February are normally the months with the highest malaria incidence [61], so a rise in case numbers may normally be expected. However, neither the district authorities nor the Epidemiology Unit of the Ministry of Health have reported any malaria cases from the affected areas for 30 December 2004 –13 January 2005, based on the spot checks performed and the review of available health information [63]. Asymptomatic infections of *P. falciparum* and *P. vivax* and dormant stages of *P. vivax* normally provide the parasite reservoir for bridging periods of low seasonal transmission (with unsuitable conditions for mosquito vectors). Under the present policy of administering primaquine in addition to chloroquine (see section on diagnosis and treatment), the reservoir of dormant stages of *P. vivax* will be low and this will delay a possible outbreak. It must be emphasized that the low level of malaria transmission in

the recent past does not guarantee that localized or even island wide epidemics will not occur. In the past, even after periods of very low levels of malaria transmission, outbreaks have occurred, often due to constraints placed on the public health system, by unusual rainfall patterns or by yet unexplained factors.
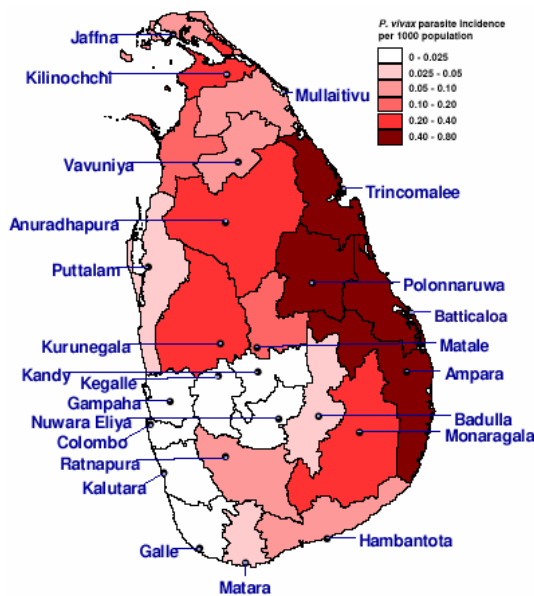
Figure 9 - Parasite incidence of *Plasmodium vivax*

Map of the districts of Sri Lanka with *P. vivax* malaria cases per 1000 population over the period January – October 2004.
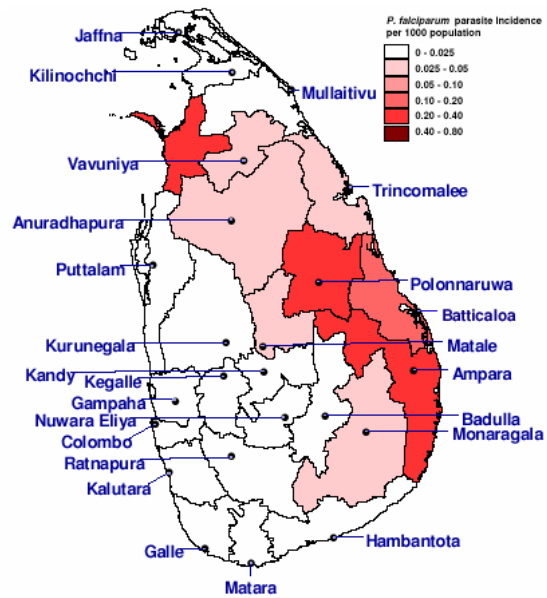
Figure 10 - Parasite incidence of *Plasmodium falciparum*

Map of the districts of Sri Lanka with *P. falciparum* malaria cases and mixed infections of both *P. vivax* and *P. falciparum* per 1000 population over the period January – October 2004.

Capacity of health care services and disease surveillance

An important factor to consider in the current situation is the capacity of the existing health care service. Following the tsunami the Sri Lanka Ministry of Health reported 22 hospitals and nine administrative buildings damaged or completely destroyed, mostly in Ampara and Trincomalee districts [64]. It has been reported that at least 40 doctors and hundreds of other medical staff have died as a consequence of the tsunami and a much higher number injured or in other ways affected by the disaster [65]. However, both the central government departments and organizations in the field report sufficient medical staff. Even in the conflict affected areas in the north and east, the AMC has been able to monitor malaria and react timely with control

measures to outbreaks since the peace process started in 2002. Also, the AMC has long standing experience with mobile clinics for malaria detection and treatment in remote areas. Lack of co-ordination among the many government departments, international aid agencies, non-governmental organizations and private individuals involved in the first phase of the emergency continues to be an important issue weeks into the disaster. According to the Ministry of Health media reports, more than 600 foreign doctors are now working in the affected areas, but few, if any, are registered with the Sri Lanka Medical Council or other relevant authorities [66]. With doctors from many countries, language barriers are also a perceived problem.

In some places, central stocks of medical supplies were destroyed, including the Regional Medical Supply Division in the Ampara District. However, sufficient drugs have been imported during the days and weeks following the disaster. The World Health Organization has drawn up plans for antimalarials, insecticides and spray equipment to be made available on request. Although the increased capacity at the district and provincial levels has improved the co-ordination, a risk remains that local needs for health care are not adequately covered in spite of the availability of significant resources. In some parts of the island, especially areas in the east, affected both by the destruction caused by the tsunami and by exceptionally heavy rainfall in the weeks following, distribution of drugs has been problematic and this has left certain communities vulnerable.

Whereas the overall capacity to provide treatment and routine malaria control activities, in general, has not been severely hampered, the routine health information system will have been constrained by the large number of autonomous health camps set up, and their lack of integration with the established surveillance system. It is essential to establish a system for monitoring malaria in the affected areas. Many people are moving back to their old place of residence trying to rebuild livelihoods and it will be essential for the public health authorities to keep contact with these communities to prevent an increase in malaria going unnoticed.

Diagnosis, treatment and drug resistance

In Sri Lanka, microscopy on blood smears or use of rapid diagnostic test kits have been the standard diagnostic procedure, and precedes the prescription of drugs to the patient. In the current situation, with the many small health clinics established within emergency camps, it is likely that the use of rapid diagnostic kits would be the more feasible means of confirmation. The first line drugs recommended for malaria treatment in Sri Lanka is still a chloroquine and primaquine (PQ) combination for cases of *P. vivax* as well as *P. falciparum* infection. Primaquine is not administered to children below one year, and those with known G-6PD enzyme deficiency, and for pregnant mothers.

So far, there have been no reports of chloroquine-resistant *P. vivax* infections in Sri Lanka. The first chloroquine-resistant *P. falciparum* case was reported in 1984 [67]. Up to 62% *in vivo* chloroquine resistance has been recorded in malarious areas [30,34,61,68]. For chloroquine resistant cases of *P. falciparum* the government recommended drug is sulphadoxine-pyrimethamine (SP). However, SP is not recommended for the last trimester of pregnancy, first six weeks of lactation and for children below two months of age. The first SP-resistant case of *P. falciparum* was reported in 1992 in Polonnaruwa district. Up to 1999, five to six cases have been reported by the AMC. More recently (January – June 2002), SP resistant *P. falciparum* has been documented in the Northern Province [68]. For SP resistant cases quinine is recommended, but only as an in-patient treatment.

In the current emergency situation, with many (foreign) doctors working autonomously, the diagnosis and treatment practices may depart from the established government guidelines and new antimalarials are also likely to be brought in. Moreover, the current practice of restricting SP to government hospitals will be difficult to enforce. Similarly, introduction of low quality and obsolete drugs will be difficult to counter at community level at the current stage of supervisory capacity and co-ordination level. Drugs have been reported stolen from warehouses, allegedly finding their way to private trade establishments [69]. Overall, it is crucial that the development of drug resistance is monitored closely and inappropriate drugs are actively phased out of the market to avoid later complications in case management.

Environmental changes and vector breeding

The seawater brought inland by the tsunami has mixed with monsoon rainwater to form puddles of varying salinity. Also, thousands of muddy surface water puddles have been created as a result of destruction and rehabilitation activities that are already underway. The brackish puddles are expected to favour the breeding of *Anopheles subpictus* sibling species B, which is a well-known coastal breeding species in Sri Lanka. However, it has not been directly incriminated as a field vector in Sri Lanka, despite its susceptibility to *P. falciparum* [70]. Nevertheless, Abhayawardana et al. [71] found peak malaria transmission in coastal areas of Puttalam in the presence of *An. subpictus* sibling species B and the complete absence of *Anopheles culicifacies* (the main malaria vector in Sri Lanka), and suggested that this *An. subpictus* sibling may have a role in transmission. It is noteworthy, that freshwater *An. subpictus* (which is now known to consist of a mixture of species A, C and D), which breeds in muddy rain fed puddles, has been consistently incriminated in malaria transmission in many inland areas of Sri Lanka [2]. Another species that is likely to breed prolifically in muddy rain-fed pools is *Anopheles vagus*. This species has been linked as a vector responsible for a malaria outbreak in southern Sri Lanka [2,72]. On present evidence, neither *An. subpictus* nor *An. vagus*, are likely to cause major malaria epidemics but could, at high density, be responsible for focal outbreaks that need quick action. Thus, it is important that an entomological monitoring programme be set up in the period leading up to and during the south west monsoon that is expected during May – June 2005 in the tsunami affected western and southern Sri Lanka. It should be noted that the infamous Asian brackish water breeding malaria vector *Anopheles sundaicus*, which is a threat in the tsunami-affected areas in Indonesia, Myanmar, and the Andaman and Nicobar islands [73], does not occur in Sri Lanka.

The main vector in Sri Lanka is *An. culicifacies* type E [74,75], which breeds mainly in pools formed in river and stream beds, and therefore, its density is mostly dependent on temporal and spatial variations in rainfall and river flow. *Anopheles culicifacies* also breeds in abandoned gem mining pits, agricultural wells and to a lesser extent in pools in agricultural water reservoirs [2]. It is unlikely that the rubble constituting a major part of the landscape in the affected areas creates breeding opportunities for *An. culicifacies,* unless it blocks waterways and creates pooling.

Post tsunami development activities may revive banned sand mining practices in rivers. If this happens, clear water pools created by these sand mining activities may serve as breeding sites for *An. culicifacies* [2]. Overall, it is very unlikely that the principal vector of malaria in Sri Lanka will breed prolifically in brackish water habitats or other habitats that may be created during the post tsunami reconstruction phase. Similarly, the principal dengue vector in Sri Lanka, *Aedes aegypti,* does not breed in saline water [76]. However, it may find plenty of rainwater-filled containers amidst the rubble created by the disaster for it to breed.

Vector control strategies and insecticide resistance

The Colombo based Head Office of the AMC gives the overall guidelines for island wide vector control, while each province works out a plan for control activities based on the distribution and level of malaria transmission. Several malaria vector control interventions are currently employed within the country. In all districts, residual insecticide spray activities are focused on areas where malaria transmission has been established by confirmed malaria cases. The control of anopheline larvae using mostly chemicals focuses on sites close to human habitation. Small-scale application of larvivorous fish and environmental modifications are also carried out. Since 1997, mosquito nets, which are biannually treated with insecticide, are distributed free of charge in malarious areas. Since two years, the main control effort is through these nets. Since January 2004, 80,000 nets with long lasting insecticide have been distributed. Also, nets are available for purchase from outlets in most parts of the country.

Studies in Sri Lanka over the 1990s on *An. culicifacies* and a range of potential secondary vectors such as *An. subpictus* and *An. vagus* have shown high level of resistance to either organochlorines, organophosphates or to both groups of insecticides [2,77-79]. DDT and Malathion are no longer recommended since *An. culicifacies* and *An. subpictus* has been found resistant. Currently, synthetic pyrethroids such as Cyfluthrin, Deltamethrin, Etofenprox, and Lambda-cyhalothrin are being used in the country. At present, Fenitrothion is the only organophosphate used for vector control. A study conducted by Abhayawardana from 1990 to 1992 [11,71] on *An. subpictus* found 68% and 54 % susceptibility to Malathion and Fenitrothion, respectively, for inland species (sibling species A), whereas for coastal species (primarily sibling species B) it was 100%. However, the latter was found

resistant to permethrin [71]. From several districts it was reported that, as a result of the tsunami, organisations have brought in insecticides not normally used or no longer recommended for vector control in Sri Lanka (P. Amerasinghe, personal communication). Vector resistance, in the light of the introduction of new insecticides, needs to be monitored and if necessary action should be taken.

Exposure of the affected community

The majority of the people initially affected by the disaster are still living in emergency camps or in other places close to the coast. At the time of writing, to the best of our knowledge, relatively few people have moved from areas of low or no malaria transmission to areas of high transmission. However, during the next phases, when people may be resettled in semi-permanent and later in permanent housing, communities may be relocated from areas where they have had no malaria experience to malarious areas. In these situations, the communities' capacity to cope with malaria infection will be low.

Despite distribution of nets to many camps, and intensified vector control in some areas, people in the emergency camps (schools, temples, mosques, etc.) and those returning to damaged houses are more exposed to mosquito bites than in pre-disaster housing, due to the open nature of the shelter. Additionally, most families will have lost mosquito nets or other means to protect against mosquito bites. It is more difficult to assess the protective effect of tents that have been set up in most of the semi-permanent camps established. The location of semi permanent and permanent settlements may have a significant effect on the risk of infection. Epidemiological studies from other parts of Sri Lanka have shown that people living within 750 m of a stream with *An. culicifacies* breeding, were at significantly higher risk for malaria than people living further away [37].

Conclusions

This paper provides maps of both *P. vivax* and *P. falciparum* malaria incidence distribution on the island of Sri Lanka at district resolution in the 10 months preceding the tsunami, and an analysis of monthly malaria incidence in the country since January 1995. The malaria incidence was historically low, which implies a limited parasite reservoir in the human population. In spite of the fact that the months of December and January are normally the peak period for transmission, given the

transmission level in the months leading up to the disaster, the risk of a large-scale outbreak seems to be limited. However, the low transmission levels over the past years may also have made people less alert to possible outbreaks, and the population would have less protective immunity towards the disease. The environmental changes resulting from the tsunami do not create particular opportunities for breeding of the principal malaria vector *An. culicifacies* but potential does exist for less important species such as *An. subpictus* and *An. vagus*. People living in emergency camps or returning to pre-disaster areas of residence are at higher risk of mosquito bites than normal. In spite of the emergency, the capacity of the public health authorities to perform malaria preventive and curative interventions remains high and essential supplies and staff capacity is not a problem. However, co-ordination of assistance and maintaining a strong surveillance system remain significant areas of concern. Increased attention to the establishment of a monitoring system including both parasitological and entomological parameters is recommended. Likewise, the large inflow of donated drugs and insecticides outside government control will potentially have long term implications on malaria control and case management, and especially the quality of administered drugs and the development of drug resistance requires careful monitoring.

Authors' contributions

GNLG collected the malaria data. OJTB checked the data, calculated incidence, made the maps. FK, FPA and PHA performed the focused literature review. All authors helped write, read and approved the final manuscript.

Acknowledgements

Chapter 4 – Malaria in Sri Lanka: one year post-tsunami

*This chapter was published in the Malaria Journal 2006, **5:**42*

Olivier JT Briët[*1], Gawrie NL Galappaththy[2], Priyanie H Amerasinghe[1] and Flemming Konradsen[3]

1 International Water Management Institute, P.O. Box 2075, Colombo, Sri Lanka, 2 Anti Malaria Campaign Head Office, Colombo, Sri Lanka and 3 Department of International Health, University of Copenhagen, Denmark

* Corresponding author

Abstract

One year ago, the authors of this article reported in this journal on the malaria situation in Sri Lanka prior to the tsunami that hit on 26 December 2004, and estimated the likelihood of a post-tsunami malaria outbreak to be low. Malaria incidence has decreased in 2005 as compared to 2004 in most districts, including the ones that were hit hardest by the tsunami. The malaria incidence (aggregated for the whole country) in 2005 followed the downward trend that started in 2000. However, surveillance was somewhat affected by the tsunami in some coastal areas and the actual incidence in these areas may have been higher than recorded, although there were no indications of this and it is unlikely to have affected the overall trend significantly. The focus of national and international post tsunami malaria control efforts was supply of antimalarials, distribution of impregnated mosquito nets and increased monitoring of the affected area. Internationally donated antimalarials were either redundant or did not comply with national drug policy. However, few seem to have entered circulation outside government control. Despite distribution of mosquito nets, still a large population is relatively exposed to mosquito bites due to inadequate housing. There were no indications of increased malaria vector abundance. Overall it is concluded that the tsunami has not negatively influenced the malaria situation in Sri Lanka.

Introduction

One year ago, the authors of this article reported in this journal [80] on the malaria situation in Sri Lanka prior to the tsunami that hit on 26 December 2004, and estimated the likelihood of a post-tsunami malaria outbreak. Here they report on changes in malaria incidence recorded in the Anti Malaria Campaign (AMC) since the tsunami. Also, they discuss the control measures taken in response to the tsunami by the Anti Malaria Campaign (additional file 1), international donors and non-governmental organizations (NGOs) (additional file 2), and effects on the surveillance system.
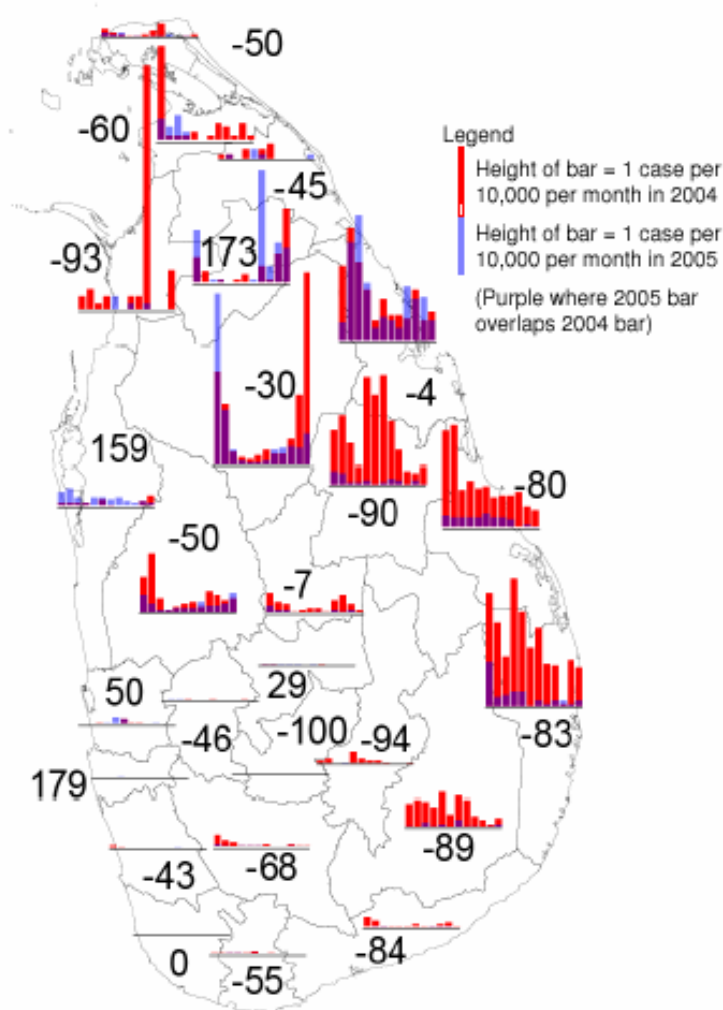
Figure 11 - Parasite incidence by district pre- and post-tsunami.

Monthly parasite incidence of *P. falciparum* and *P. vivax* malaria combined in 2004 (red bars) and 2005 up to December (blue bars). The bar on the far left represents January. Overlapping bars colour purple. The height of the bars in the legend represents 1 case per 10,000 persons per month. Numbers indicate the percentage change per district of January – December 2005 as compared to January – December 2004.

Malaria incidence pre and post tsunami

Figure 11 shows the monthly malaria incidence January 2004 – December 2005 at district level, and the percent difference between the successive years. On the north, east and south coasts, most of the districts show a strong decline (except Trincomalee [-4%]), whereas the malaria incidence increased in the districts of Colombo, Gampaha and Puttalam on the west coast, which were the least tsunami affected coastal districts of the country. A possible reason for the slow decline of malaria in Trincomalee, as compared to other coastal districts in the east, is that particularly in Trincomalee there was increased civil unrest and political protests in 2005 as compared to 2004, hindering the control efforts of the regional malaria officer (RMO). There is no indication that the physical impact of the tsunami has affected the malaria case load in Trincomalee District as 85% of the cases were reported from Trincomalee Medical Officer of Health (MOH) administrative area, (representing about 28% of the population in the District), which was relatively unaffected by the tsunami in terms of buildings destroyed and people killed or missing [81] as compared to other coastal MOH administrative areas in Trincomalee district. Only eighteen cases (six percent of the total cases in Trincomalee District) occurred in Kinnyia MOH among tsunami-displaced people. Note that despite the "dramatic" increases in malaria incidence in some districts on the west coast, the number of cases in 2005 was low, and these increases could have arisen from incidental mini-outbreaks.

Table 4 - Annual percentage of growth in malaria incidence over the years 2001 to 2004 as compared to the years 2004 – 2005, absolute case numbers in 2005 and population projection in districts in Sri Lanka, and their geographic position.

District	Position	Annual percent growth of incidence 2001-2004	Annual percent growth of incidence 2004-2005	Difference	Absolute cases 2005	Projected population mid 2005
Ampara	East coast	-18	-83	-65	126	612696
Anuradhapura	Inland	-50	-30	20	448	780900
Badulla	Inland	-65	-94	-30	3	829245
Batticaloa	East coast	-59	-80	-22	84	540535
Colombo	West coast	-61	179	240	17	2382298
Galle	South coast	NA	NA	NA	0	1028690
Gampaha	West coast	-63	50	113	56	2112382
Hambantota	South coast	-65	-84	-18	6	536952
Jaffna	North coast	-70	-50	19	24	653466
Kalutara	West coast	-52	-43	8	4	1092711
Kandy	Inland	-59	29	88	17	1345092
Kegalle	Inland	-65	-46	19	6	798845
Kilinochchi	West coast	-89	-60	29	16	147603
Kurunegala	Inland	-61	-50	11	258	1503423
Mannar	West coast	-37	-93	-56	4	100181
Matale	Inland	-50	-77	-27	18	462498
Matara	South coast	-47	-55	-9	10	799400
Moneragala	Inland	-64	-89	-25	17	413301
Mullaitivu	East coast	-88	-43	45	5	122942
Nuwara Eliya	Inland	-55	-100	-45	0	728166
Polonnaruwa	Inland	-41	-90	-49	37	378379
Puttalam	West coast	-69	159	228	102	738475
Ratnapura	Inland	-74	-68	6	22	1068896
Trincomalee	East coast	-37	-4	33	286	355573
Vavuniya	Inland	-73	45	118	62	155650

NA*: In Galle District, only 2 malaria cases were reported in 2001 and none thereafter.
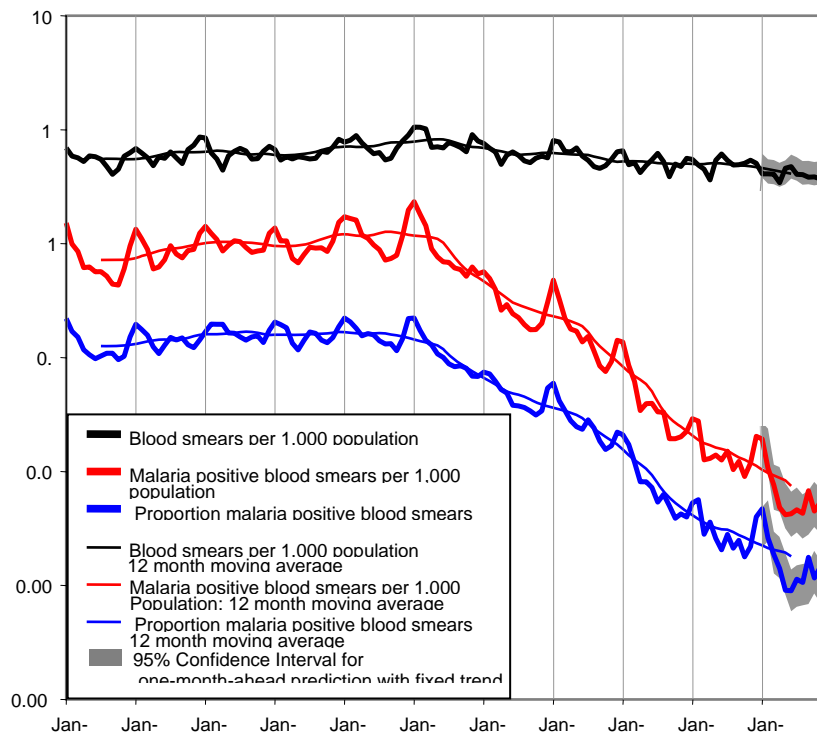
Figure 12 - Monthly parasite and blood smear examination incidence patterns.

Monthly parasite incidence patterns of *P. falciparum* and *P. vivax* malaria combined per 1000 population (thick red line) and 12 month moving average (thin red line), blood smears examined per 1000 population (thick black line) and 12 month moving average (thin black line), and proportion of blood smears positive for malaria (thick blue line) and 12 month moving average (thin blue line) from January 1995 to December 2005 in Sri Lanka. Ninety-five % Confidence areas are indicated for one-month ahead prediction of time series for January - December 2005, using the fixed trends as found in the period April 2000 – December 2004 (in grey).

Although with considerable spatial variation (Table 4), the malaria incidence in Sri Lanka shows a general downward trend since about April 2000 (Figure 12). This makes it necessary to assess whether this trend has continued after the tsunami, or whether there has been a change in trend since the disaster. A possible explanation for the downward trend pre-tsunami is the decrease in armed conflict that preceded the signing of the memorandum of understanding on the Permanent Cessation of Hostilities in February 2002. The Anti Malaria Campaign was one of the few institutions that had access to uncleared areas during the conflict, but was limited in its operation by hostilities. It should be noted that malaria transmission in Sri Lanka has always fluctuated over the years in response to major changes in control strategies

51

and efforts, climate variations, or due to factors not yet fully established. Between April 2000 and December 2004, the malaria positive blood smear incidence in the whole country has decreased exponentially by 8 percent per month (the 13 month moving average [with values for months at extremes given half weight] plotted in Figure 12 is approximately a straight line on a logarithmic scale). Over the same period, the surveillance effort (the blood smears examined per population) has also decreased exponentially, but only by 0.8% per month. By December 2004, the ratio of positive slides to the number of blood slides examined was 1:250 (400 / 99694). The observed monthly incidence since the tsunami, over the period January – December 2005 is not significantly different (alpha = 0.05) from the one-month-ahead prediction applying the fixed trend and using a first order auto regressive model with first order seasonal component. Therefore, at the country level, there is no evidence that the tsunami affected the incidence of malaria.

Surveillance issues

Many of the tsunami-affected areas are also affected by the ethnic conflict, and this undoubtedly influences the quality of incidence records negatively in these areas. This, and spatial differences in other surveillance aspects [61] and transmission make it logical to assess the impact of the tsunami disaster and response by comparing the malaria situation in affected areas post disaster with the situation preceding the disaster, rather than comparing affected with unaffected areas. However, surveillance was not quite the same in affected areas post tsunami. Mobile clinics have visited camps with displaced people and performed active case detection (patients with fever were encouraged to test for malaria), and results were not taken into account in the surveillance database. Furthermore, approximately 6000 malaria rapid diagnostic kits have been supplied by UN agencies [82]. These kits are used by AMC Regional Malaria Officers in remote areas if no microscopists are available. Results from these tests were not included in the surveillance statistics in 2005 (but will be from January 2006 onwards). At the country scale, 1000 kits per month is relatively insignificant to the 100,000 slides that were on average examined (routinely) monthly over 2005, but this may have underestimated figures from tsunami affected areas specifically. However, the number of cases detected but not reported by these methods can be characterized as "few", although precise figures are not available (GNLG, personal communication). Overall, the surveillance effort (blood films examined) over the

period January –December 2005, as per the surveillance database (Figure 12) was not significantly different from predicted. Therefore, at country level, there is no indication that the tsunami has affected surveillance capacity, either negatively or positively after the international aid effort.

Tsunami medical aid and interference with country malaria treatment policy

Over 100,000 anti-malarial tablets were supplied by UN agencies after the tsunami [82]. These are chloroquine, proguanil and sulphadoxine / pyrimethamine (SP) tablets, which are in line with government policy. However, given the current low endemicity level of malaria (less than 4,000 cases were reported over the year 2004, and 1,628 cases over 2005), and the fact that government drug warehouses were well stored (50,000 tablets per district) prior to the tsunami, most of these donated tablets is likely to expire. Although stocks of some other medications have been reported "lost" from warehouses, this has not been the case for antimalarials. Some NGOs brought small quantities of artesunate-based tablets (currently not in line with government policy) into the country, and were treating fever cases in camps the first week after the disaster until being asked by the AMC (RMOs visited camps when possible) to refer cases to government medical institutions or otherwise treat with government approved medication after confirmation of disease. A large shipment of unsolicited artesunate-based tablets arrived at Colombo port but was not cleared by customs after AMC orders (GNLG, personal communication).

Vector control and personal protection since the tsunami

Half a year post tsunami, over 500,000 people were registered as tsunami-displaced in welfare centres or staying with friends and relatives [82], of which then almost 10,000 families were still staying in tents. The latest government estimates are that Sri Lanka is 21 percent of the way to its overall housing goal. So far, 7,461 new homes have been built, while homeowners have repaired another 13,737 homes. These statistics are from the government's Reconstruction and Development Agency, which is coordinating the tsunami recovery. Therefore, by government estimates, several hundred thousand Sri Lankans are still without permanent homes. Some 33,000 families, or at least 150,000 people, remain in transitional shelters. Others are living temporarily with relatives or friends [83]. Although international and national NGOs have contributed to malaria vector control by distributing over 100,000 insecticide

treated nets (Additional file 2), many of the people in the camps have no way to use a mosquito net because the tents or barracks are too crowded [84], or they have difficulty placing the nets properly (particularly in tents). Also, lack of space may force displaced people to spend lots of time outside and therefore exposed. Although in tsunami affected Aceh (Indonesia), insecticide-impregnated plastic sheeting for refugee tents and temporary houses were used, these have not been reported in Sri Lanka. International NGOs have also contributed to malaria vector control by supplying fogging machines and they have helped with the development of disease awareness campaigns. After incidental reports of NGOs spraying insecticide, the government has issued technical guidelines on the application of insecticides for control of vector-borne diseases in tsunami-affected areas in January 2005. In the immediate post tsunami period, there was a programme of chlorination of wells and regular spraying of larvicides and fogging around settlement camps. Also, in some camps, RMOs have carried out indoor residual spraying. In this context, it is important to note that for *Anopheles culicifacies* and *Anopheles subpictus* A, widespread resistance to Malathion has been found and, maybe more importantly, these two species also have developed some resistance to pyrethroid insecticides now important in the operations of malaria control [85].

Population movement

In South India, the potential migration in and out of the tsunami hit areas was initially seen as a potential risk of introducing malaria into areas with low prevalence [86]. However, a situation analysis three months after the disaster found only limited population movements in South India but highlighted the problems of increased migration into the area by e.g. fishermen and pilgrims [87]. Although in Sri Lanka, (still) many people are registered as Internally Displaced Persons (IDPs) as result of the tsunami, most of these are hosted by friends and family or housed in camps in the vicinity of their original residences. Figures on inter-district displacement as result of the tsunami are not available, but it can be characterized as low (International Organization for Migration Sri Lanka, UNHCR Sri Lanka, Humanitarian Information Centre for Sri Lanka, personal communications).

Vector ecology

Densities of the principal malaria vector *An. culicifacies* and secondary vector *An. subpictus* have declined in the whole country, possibly due to a gradual switch from less effective organophosphates to pyrethroids since 2001, but it is also possible that indoor resting of vectors has been affected by the decline in use of thatched roofs (with *cadjan* or *palmyrah* leaves) with socio-economic development (The districts in the north and east have profited less from this development). Overall, there was no indication of increased malaria vector breeding in tsunami affected areas (R. R. Abeyasinghe, personal communication). In the tsunami affected southern District of Hambantotha, not a single *An. culicifacies* specimen was found by the AMC entomological team in the normally high transmission month January. In a small study in tsunami-affected areas in the east coast, it was found that anophelines (*An. culicifacies* in wells, *An. subpictus* in pools after rains, *Anopheles varuna* and *Anopheles vagus* in rice fields) were breeding in all types of habitats, albeit in small numbers. Larval populations were affected by rainfall and cleaning of wells [88], as well as by chlorination and regular spraying of larvicides. On average, the coastal belt affected by salt water intrusion was about 1 km from the shore, and the salt water deposited by the tsunami dried up quickly and did not sustain the breeding populations of mosquitoes. The east of the island (especially Batticaloa and Ampara Districts) received substantial rainfall within a couple of weeks after the tsunami, resulting in some flooding. This fresh water receded within a few days washing away any traces of salt water. An entomological study undertaken post tsunami along the coast of Tamil Nadu (India) found that the urban malaria vector *Anopheles stephensi* was introduced into the area from nearby towns and became the predominant species only four weeks after the wave hit the coastline. Likewise, in India, the rural malaria vector *An. culicifacies* seemed to adapt to higher levels of salinity than what had previously been reported, but the increasing temperatures dried up the breeding sites created in the debris left behind by the tsunami by April-May [87]. On the Andaman and Nicobar Islands, the most important vector found breeding in tsunami affected areas was the brackish water mosquito *Anopheles sundaicus* [89]. In Sri Lanka, neither *An. sundaicus* nor *An. stephensi* are found.

Conclusions

Despite initial warnings from some international health agencies, malaria in Sri Lanka did not increase after the tsunami. The effect of the tsunami disaster on malaria can not be accurately assessed in the presence of control measures. However, it appears that measures taken by the local health authorities in collaboration with NGOs were sufficient to prevent possible outbreaks. In addition, the ecological impact of the tsunami was not conducive to malaria vector breeding. However, still a large part of the population may be exposed to mosquito bites, despite distribution of impregnated bed nets. It is unfortunate that a large quantity of antimalarials may go to waste due to over abundant stocks, or incompatibility with the national drug policy. In emergency situations, donors and NGOs are urged to contact local health authorities for coordination.

Authors' contributions

OJTB Initiated the paper and did the analysis; GNLG supervised collection of malaria case data, and provided critical information on AMC policy; FK performed literature analysis; PHA supervised the mosquito breeding study in the tsunami affected area in the east. All authors helped write, read and approved of the final manuscript.

Acknowledgements

List of additional files to this chapter (printed in Appendix)

Additional file 1

Short/Medium Term Plan for prevention and control of possible malaria outbreaks in tsunami affected areas of Sri Lanka.

Additional file 2

Non-exhaustive list of reported antimalarial support by non-governmental organizations.

Addendum to Chapter 4 – Analysis of pre and post tsunami trends in malaria incidence time series

Introduction

The availability of more recent malaria case data, and the methodology developed in Chapter 8, allowed for a more rigorous statistical analysis of the question whether the tsunami has had an impact on all-island malaria incidence. In particular, it was investigated whether post-tsunami the slope of the (downward) trend changed.

Material and Methods

Data

The time series of whole island malaria cases and population used were the same as presented in Chapter 4, with the exception that the time series were extended up to and including December 2006 (Figure 13).
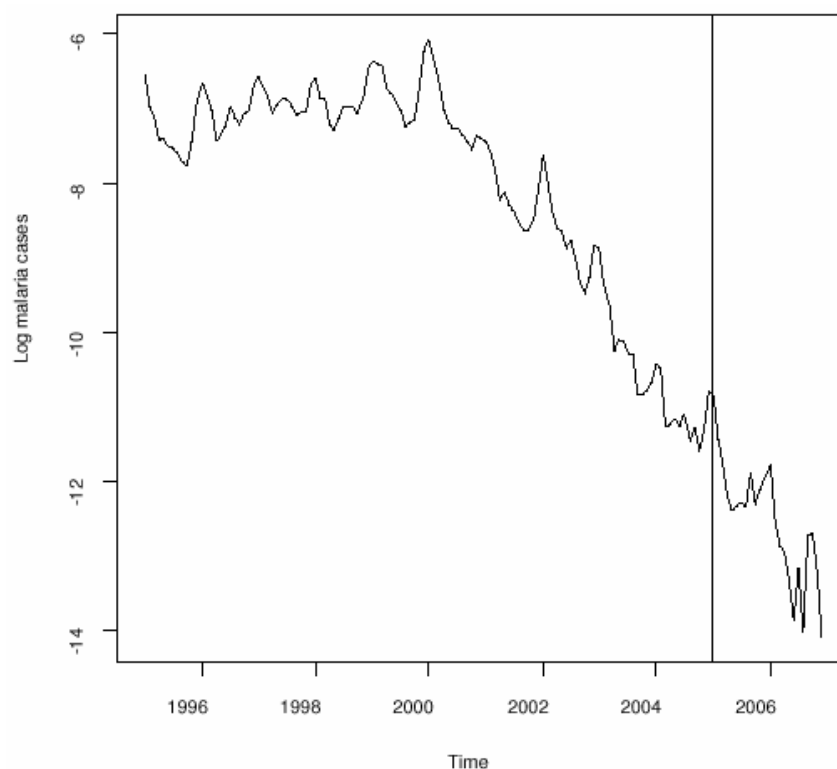
Figure 13 - Malaria incidence in Sri Lanka 1995 – 2006

Logarithmically transformed monthly malaria incidence aggregated for the whole country of Sri Lanka from 1995 to 2006. The vertical line indicates the day when the tsunami hit Sri Lanka.

Statistical methods

The statistical methods are similar to those employed in Chapter 8. Unlike in Chapter 8, where malaria time series were presumed to have a stochastic trend, in the present analysis, malaria incidence data were presumed to follow a deterministic trend. The data were thus not differenced, but an intercept and a piecewise linear trend was fitted to the data [90,91]. A piecewise linear trend contains intervals where the trend line is locally linear, but there are change points where the slope may change abruptly.

The piecewise linear trend was modelled by including a variable for each interval for which the slope was modelled separately. *E.g.* a model allowing for a slope up to a certain change point, and a different slope thereafter thus contained two trend variables. Each trend variable $x_{k,t}$ for the local interval k followed the function:

$$x_{k,t} = \begin{cases} 0 & \text{for } t < a_k \\ t - a_k + 1 & \text{for } a_k \le t \le b_k \\ b_k - a_k + 1 & \text{for } b_k < t \end{cases}$$

with a_k the start month and b_k the end month of the local interval k, where $a_1 = 1$ and $a_k = b_{k-1} + 1$. The sum of all the trend variables resulted in the function

$$\sum_{k=1}^{k=K} x_{k,t} = t,$$ with K the total number of slope-intervals.

The incidence pattern (Figure 13) shows a clear seasonality, and a change in slope at the beginning of the year 2000 is apparent. A change in slope after December 2004 is not evident, although the peak at the end of 2005 and 2006 might suggest a slightly less strong negative slope.

In an exploratory analysis, logarithmically transformed malaria incidence was assumed to be Gaussian distributed. In subsequent analysis, the malaria cases were assumed to be negative binomially distributed, and population was used as an offset variable.

In Gaussian (S)ARMA analysis with $K = 2$, all combinations of parameters p, q, P, and Q were tested with p, q, P, Q $\in \{0,1,2\}$, and $b_1 \in \{55,\ldots, 66\}$ (corresponding to the interval July 1999 – June 2000). In order to model seasonality, two sub-models were considered: a model with deterministic seasonality through second order harmonics, and a model with seasonal random effects. Akaike's information criterion

(AIC) was used to determine the values of the parameters giving the best fit for each sub model. The statistical package "R" was used for the analysis.

The selected (S)ARMA models were consequently modelled in a Bayesian framework with a negative binomial distribution, using the software package WinBUGS (see Chapter 8). The models were then fitted with an additional interval, allowing the slope to be different after the tsunami event, thus in models with $k = 3$ and $b_2 = 120$ (corresponding to December 2004). The deviance information criterion (DIC) was used to determine whether allowing the slope for January 2005 – December 2006 to differ from the slope from 2000 – December 2004 improved the model.

Results and conclusion

The best SARMA model was the model SARMA(p=1, q=0, P=1, Q=1), with a change point between February and March 2000 (Henceforth Model 1a). The best ARMA model with second order harmonics was the model ARMA(p=1, q=0), with a change point between March and April 2000 (Henceforth Model 2a).

In the Bayesian negative binomial models, the addition of a change point in slope between December 2004 and January 2005 to Model 1a (resulting in Model 1b) or Model 2a (resulting in Model 2b) did not yield a lower DIC, in either model (Table 5).

Therefore, it can be concluded that the tsunami did not cause a change in slope in malaria incidence trends.

Table 5 - Results from negative binomial models

Parameter/Criterion	Model 1a		Model 1b		Model 2a		Model 2b	
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.
$\beta_{intercept}$	-6.33	0.985	-6.16	0.512	-7.30	0.114	-7.30	0.112
ϕ_1	0.67	0.071	0.68	0.071	0.60	0.069	0.61	0.067
ϕ_1^*	0.93	0.049	0.92	0.054				
θ_1^*	-0.66	0.144	-0.63	0.151				
$\beta_{\sin(2\pi t/12)}$					0.14	0.038	0.14	0.039
$\beta_{\cos(2\pi t/12)}$					0.24	0.037	0.24	0.038
$\beta_{\sin(2\pi t/6)}$					0.16	0.025	0.16	0.024
$\beta_{\cos(2\pi t/6)}$					0.08	0.024	0.08	0.024
β_{x_1}	0.01	0.007	0.01	0.006	0.01	0.003	0.01	0.002
β_{x_2}	-0.08	0.004	-0.09	0.005	-0.08	0.002	-0.08	0.003
β_{x_3}			-0.10	0.023			-0.08	0.010
r	30.11	4.736	29.11	4.554	34.15	4.828	33.93	0.067
DIC	2017.08		2021.31		2227.87		2229.85	

Legend: x_1 = local linear trend for the interval from January 1995 up to and including February 2000 (model 1ab) or March 2000 (model 2ab); x_2 = local linear trend for the interval from March 2000 (model 1ab) or from April 2000 (model 2ab) up to and including December 2006 (model 1a and model 2a) or December 2004 (model 1b and model 2b); x_3 local linear trend for the interval from January 2005 to December 2006.

Chapter 5 – Malaria seasonality and rainfall seasonality in Sri Lanka are correlated in space

Olivier JT Briët[*1,2] , Penelope Vounatsou[2] , Priyanie H. Amerasinghe[3]

1 International Water Management Institute, P.O. Box 2075, Colombo, Sri Lanka, Telephone: +94112787404, Fax: +94112786854

2 Swiss Tropical Institute, Socinstrasse 57, P.O. Box CH-4002, Basel, Switzerland

3 International Water Management Institute Sub Regional Office for South Asia, c/o ICRISAT, Patancheru, AP 502 324, Andhra Pradesh, India

* Corresponding author

Abstract

Correlation in space between seasonality of malaria and seasonality of rainfall was studied in Sri Lanka. A simple seasonality index was developed by making use of the bimodal seasonality of both malaria and rainfall. The malaria seasonality index was regressed against the rainfall seasonality index taking spatial autocorrelation into account. Despite the presence of spatial autocorrelation, the coefficient for the rainfall seasonality index in explaining the malaria seasonality index was found to be significant. The results suggest that rainfall is an important driver of malaria seasonality.

Keywords: malaria, rain, seasonality, Sri Lanka
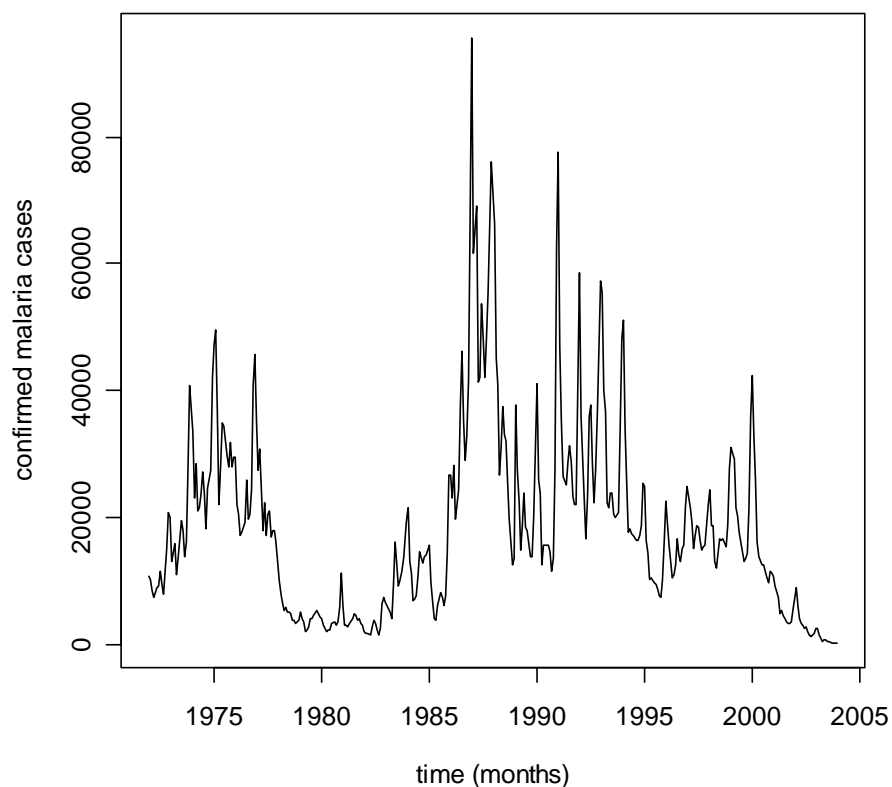
Figure 14 - Monthly confirmed malaria cases in Sri Lanka

Monthly confirmed malaria cases in Sri Lanka over the period January 1972 to December 2003.

Introduction

In the field of malaria transmission and prediction modelling there is not only an interest in risk *per se* but also an increasing interest in identifying the seasonality of malaria over larger geographical areas [52,92-95]. Malaria case time-series in Sri Lanka show both strong long-term fluctuations and seasonality (Figure 14). The long-term fluctuations are generally attributed to the impact of malaria control strategies and the development of insecticide resistance. Seasonality, on the other hand, is generally attributed to climatic factors, in particular rainfall which provides the breeding habitats for the malaria vector mosquitoes and sustains the aquatic, immature stages of their life cycle. The optimum amount and frequency of rainfall depend on the physical nature of the breeding site and the requirements of the vector in question. It should be remembered that rainfall may also have an indirect impact on mosquito breeding, e.g. when it occurs upstream or when rain water is transported for irrigation purposes. Hence, rainfall impinges on mosquito population dynamics in a rather complex way. For example, a large amount of rain within a short period of time may wash away aquatic stages as well as adults, while continuous, low-volume rainfall may not be optimal for colonizing mosquito species that require temporary breeding sites. Although malaria-case time-series (from which long term non-seasonal trends were removed) and rainfall time-series appear to have strong cross-correlations, a large part of these correlations can be explained by both series being cyclical with a similar periodicity [96]. Many biological processes follow annual cycles and high cross-correlations do not necessarily infer a causal link. However, there is spatial information that suggests that rainfall seasonality could be a driver of malaria seasonality and although the island of Sri Lanka is only 65,610 km^2, in total, it shows strong spatial variability in climate [97]. The south-western part of the country (often described as the wet zone since it receives more than 1900 mm of rain annually) is affected by two periods of monsoon rains with peaks in May and October, whereas rainfall peaks in November/December with a very minor, almost imperceptible peak in April (Figure 15) in the so called dry zone in the north-east which receives less than 1900 mm rain annually. Corresponding to this distribution of rainfall, malaria-case time-series show a strong bimodal seasonality in the south-western part of the island, whereas the malaria time-series become more monomodal in nature towards the north and east with the second peak in the middle of the year

being much less important (Figure 16). In order to establish whether there is a correlation between rainfall seasonality and malaria seasonality over space, a regression analysis was carried out.
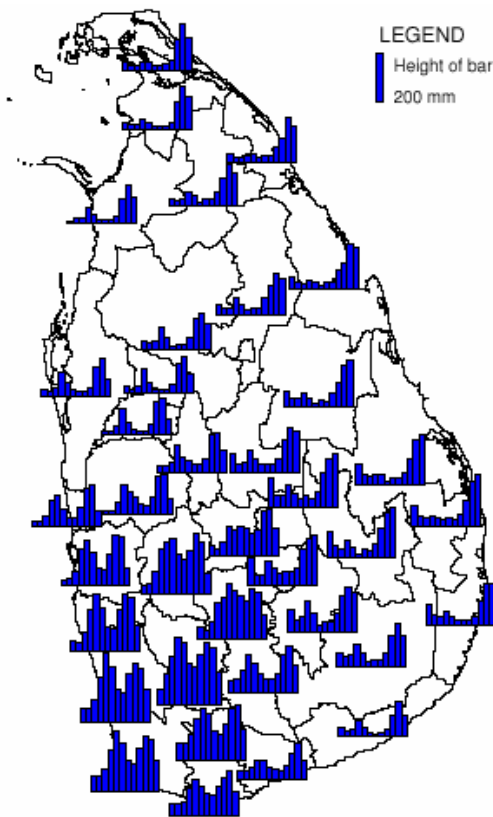
Figure 15 - Geometric mean seasonality of rainfall

Geometric mean monthly rainfall from January (bar on far left) to December (bar on far right), over the period January 1972 to December 2003, in areas of Sri Lanka. The height of the bars in the legend represents 200 mm.
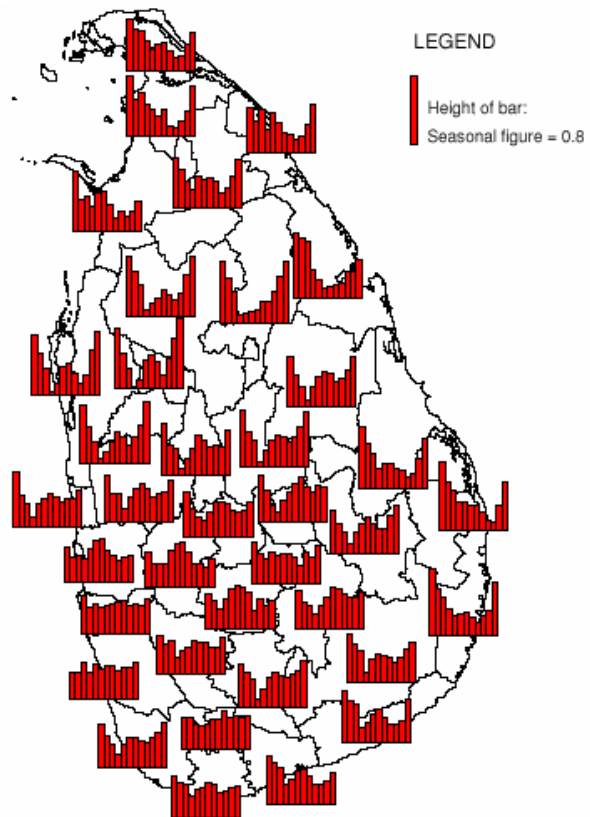
Figure 16 - Seasonal figure of malaria

Seasonal additive figure of monthly logarithmically transformed malaria cases from January (bar on far left) to December (bar on far right), over 1972 to 2003, in areas of Sri Lanka. The value of 0.6 is added to all data to make all data positive.

Materials and methods

Malaria and rainfall data

Records of the total count of blood films examined for malaria, and how many of these are positive for malaria, are reported monthly by Government health facilities and aggregated by the medical officer responsible for each so called health area (a sub-district health administrative division). This study is based on information regarding such blood film counts provided by the Anti Malaria Campaign (AMC) of

Sri Lanka for the period 1972 – 2003. For some of the records, the number of blood films examined was marked as "not received" (and therefore classified as missing). For 14.90% of the records, the value was given as zero or left blank. For the latter records (the blanks), there was ambiguity as to whether the data were missing due to problems in data recording or whether they could also be taken as zero, i.e. no patients presented themselves for examination in that particular area in that particular month. In the data cleaning procedure (see statistical methods below), 1.4% of the records were declared as not available (NA). This included the records where the place for entering the number of blood films examined was marked as "not received" (0.95% of all the records) and the records for which the number of the blood films could be classified as a lower additive outlier [98] (0.44% of all records). The data from the districts in the north and east, where data gathering and reporting was affected by the armed conflict, had the largest percentage of NA labels: Jaffna (5.4%), Mannar (26.1%), Vavuniya (8.9%), Kilinochchi (2%), Trincomalee (2%) and Ampara (5.4%). Over time, some health areas changed boundaries or split into two, e.g. in 1972 and 2003, the number of health areas was 98 and 230, respectively, and were therefore deemed unsuitable for temporal data aggregation. For the purpose of this study, health areas with variable boundaries were aggregated into larger areas corresponding to malaria data for which the catchment area did not change over the 1972 – 2003 period. Thus, the surface of Sri Lanka was divided into 37 areas (Figure 15 and Figure 16) and the "cleaned" monthly malaria-positive data were aggregated accordingly.

Precipitation records, collected by 342 stations across the island, were purchased from the Meteorological Department of Sri Lanka and monthly rainfall surfaces were created through spatial prediction using kriging. Three stations with consistently aberrant rainfall records, detected through cross-validation using kriging [99], were removed from the dataset. From each monthly rainfall surface, the average value of rainfall was extracted for each area.

Statistical methods

In a data cleaning procedure, the time series of blood film counts in health areas were logarithmically transformed to normality (after the value one was added to the data). Under the null hypothesis, each observation was assumed to be part of a seasonal autoregressive integrated moving average (SARIMA) process [100] with parameters

p=0, d=1, q=1, P=0, D=1, and Q=1. Observations were marked as additive outliers if the likelihood ratio test statistic (for an additive outlier) for the observation was below a threshold of -6 [98] and classified as NA. For each of these and other NA observations that were not at the beginning or at the end of a series, values for the number of malaria-positive blood films were estimated through a one-step-ahead SARIMA forecasting model on the original series and on the reversed series. These two estimates were then averaged. This approach has been discussed by Mwaniki and colleagues [101]. Finally, the health area data series were aggregated to the larger area (previously defined) resolution before analysis, as these spatial units remained constant over the study period, whereas for many health areas boundaries changed (within the larger area boundaries) over the study period.

For each area and for each calendar month of the year, the 34-year (the period matching the malaria data available) mean rainfall was calculated and the values logarithmically transformed. For each area, the first rainfall peak was calculated as the sum of the rainfall climatology during the calendar months March – August, and the second rainfall peak was calculated as the sum of the rainfall climatology during the calendar months September – February. The logarithmically transformed ratio of the two peaks was used as an index of rainfall seasonality. For malaria-case count time-series, a similar procedure was applied except that the long-term trends were calculated using a 13-point moving average filter with the coefficients at the extremes given half weight [90] and removed. Also, the first malaria peak was calculated as the sum of the seasonal figure during the calendar months April – September and the second malaria peak as the sum of the seasonal figure during the calendar months October – March.

The distribution of the malaria seasonality index was tested for normality using the Shapiro-Wilk test [102]. The presence of spatial autocorrelation of the malaria seasonality index among areas was tested with the Moran's I test [103].

Let y_i be the malaria seasonality index in area i, $i = 1, \ldots, 37$. It was assumed that y_i arises from a normal distribution with mean μ_i and precision parameter τ, that is $y_i \sim N(\mu_i, \tau)$. We considered $\mu_i = \beta_0 + \beta_1 x_i + \varphi_i$ where x_i measures rainfall seasonality. β_0, β_1 are regression coefficients and φ_i is an area-specific random effect taking into account the spatial correlation introduced by the spatial structure of unobserved covariates. For the φ_i's, a conditional autoregressive model with random effects,

CAR(γ, τ_s), was adopted which assumes that $\varphi_i \mid \varphi_{-i}, \tau_s, \gamma \sim N\left(\dfrac{\gamma \sum \varphi_{-i}}{n_i}, \dfrac{1}{n_i \tau_s}\right)$ where γ

is a spatial correlation parameter, n_i is the number of areas bordering area i and

$\sigma_s^2 = \dfrac{1}{\tau_s}$ measures between area variation. To facilitate model fit a Bayesian

modelling framework was used. The following prior distributions were chosen for the parameters: $\beta_0, \beta_1 \sim U(-\infty, \infty)$, τ, $\tau_s \sim Ga(0.005, 0.005)$, $\gamma \sim U(a, b)$ with limits a, b specified as described by Gelfand and Vounatsou [48]. The CAR(γ, τ_s) was considered as a prior distribution for the φ_i's. The effect of this prior distribution is to shrink the outcome variable value (in this case the malaria seasonality index) of an area to that of the local mean, where the local mean is the mean of all contiguous areas excluding the area itself. The posterior distribution of the seasonality index of an area is therefore a compromise between the prior, which is based on the seasonality index of neighbouring areas, and the data for the area. Two spatial models were fitted: a) a CAR(γ, τ_s) and b) a CAR(1, τ_s). The latter model assumes maximum spatial correlation although it does not give a proper distribution for the φ_i's [104]. The former model gives a well-defined proper distribution. Bayesian CAR models have been widely used in malaria mapping [50-52]. A non spatial model was also applied. The deviance information criterion (DIC) [53] was used to determine the best fitting model. The models were estimated using a Markov Chain Monte Carlo process using three chains, and 1,500,000 iterations (including a burn-in of 500,000 iterations), with a thinning rate of 100. Convergence was assessed by studying plots of the Gelman-Rubin convergence statistic as modified by Brooks and Gelman [54].

The regression analysis described above was repeated including zone as a regressor. Also, because there is some concern that cases in areas with historically low transmission may primarily have been acquired elsewhere (and patterns would therefore not represent local transmission) the analysis (without zone as regressor) was repeated excluding nine areas with a geometric mean annual case-load of less than 600 over the study period (all of these except Chilaw were situated in the wet zone and some were situated at high elevations): Colombo district, Kalutara district, Galle district, the northern part of Matara district, the western part of Ratnapura district, Nuwara Eliya district (comprising of two areas), eastern part of Kandy district, and Chilaw (the southern part of Puttalam district).

All data management and analysis was performed in the software environment R [56]. The Bayesian regression analysis, which was carried out in the software package "WinBUGS" [55] which can be called from R using the R package "R2WinBUGS".
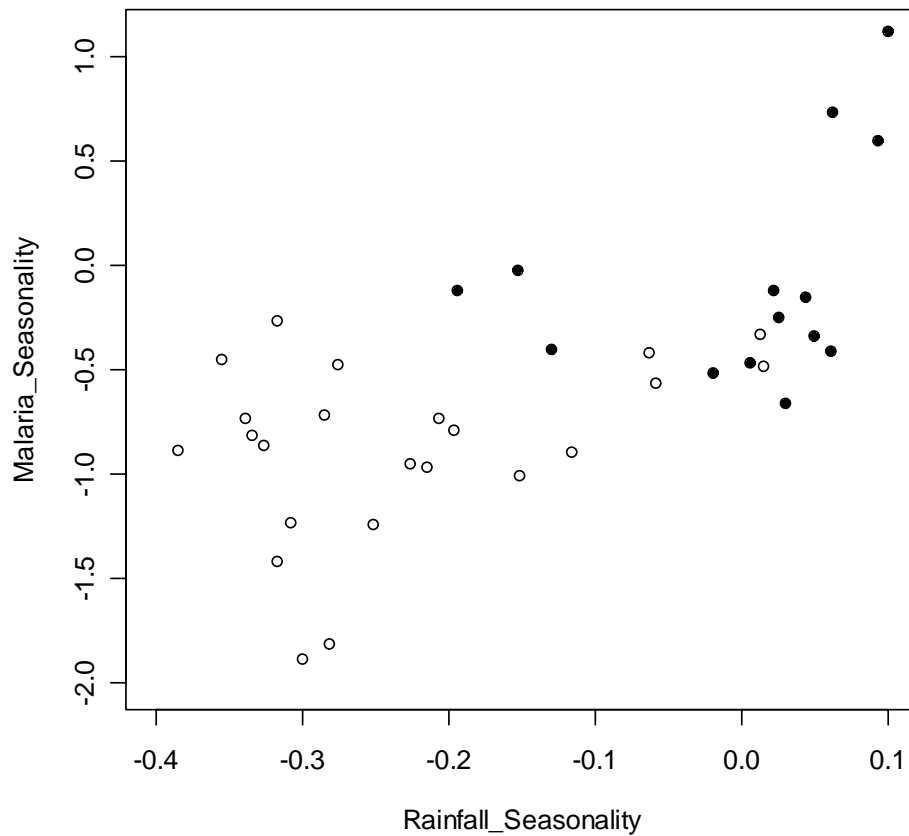
Figure 17 - Scatter plot of malaria seasonality versus rainfall seasonality

Scatter plot of the malaria seasonality index on the vertical axis against the rainfall seasonality index on the horizontal axis for areas in the wet zone receiving more than 1900 mm rainfall annually (solid circles) and in the dry zone (open circles) in Sri Lanka.

Results

There was no evidence that the malaria seasonality index, nor the residuals of any of the regression analyses, were not Gaussian distributed, according to the Shapiro-Wilk test. A scatter plot of the malaria seasonality index against the rainfall seasonality index (Figure 17) showed that there are no clear outliers. The Moran's I test for the complete dataset (n=37, I = 0.46, 95% credible interval = 0.45 – 0.52) and the reduced

dataset (n=28, I = 0.45, 95% credible interval = 0.42 – 0.54) showed that there was significant spatial autocorrelation in the malaria seasonality index, thus nearby pairs of districts had a more similar malaria seasonality index than distant pairs.

Table 6 - Results of regression of the malaria seasonality index against the rainfall seasonality index in malarious areas

Model	parameter	Mean	Standard deviation	95% Credible interval
Non spatial	β_0	-0.20	0.11	-0.41 - 0.01
	β_1	2.53	0.51	1.54 - 3.55
	τ	4.65	1.10	2.68 - 6.92
	Moran's I	0.16	0.04	0.12 - 0.27
	DIC	52.18		
CAR normal	β_0	-0.18	0.14	-0.44 - 0.09
	β_1	2.69	0.73	1.29 - 4.19
	τ	10.34	23.34	2.99 - 52.56
	τ_s	40.90	80.37	1.00 – 276.70
	Moran's I	0.11	0.11	-0.10 - 0.34
	DIC	45.47		
CAR proper	β_0	-0.19	0.13	-0.44 - 0.05
	β_1	2.61	0.56	1.53 - 3.73
	τ	10.71	32.20	2.90 - 61.91
	τ_s	53.50	93.04	0.94 - 314.20
	γ	-0.03	0.70	-1.37 - 0.96
	Moran's I	0.12	0.09	-0.09 - 0.27
	DIC	41.1		

Legend: tau = precision = 1/variance; gamma = spatial auto correlation coefficient; Deviance = posterior mean of -2*log(likelihood) ; DIC = Deviance information criterion

Estimates for the mean of the coefficients in the non-spatial, CAR(1, τ_s) and CAR(γ, τ_s) models (Table 6) were very similar. The standard deviations of the coefficients for the rainfall seasonality index and zone in explaining malaria seasonality were larger (as expected) in the spatial models as compared to the non-spatial model. The 95% credible interval of the posterior distribution of the rainfall seasonality index coefficient did not include zero. The CAR(γ, τ_s) was the best fitting model because the DIC had the smallest value. In the CAR(γ, τ_s) model with both the rainfall index

and the zone as covariate, the 95% credible interval of the posterior distribution of the rainfall coefficient was larger but still did not include zero (results not shown). It should be noted that the zone coefficient was highly significant in explaining the rainfall seasonality index (p<0.00001), thus collinearity may have played an important role. There was no evidence for zone-rainfall seasonality interaction (analysis not shown). After accounting for the effect of the rainfall seasonality (and the zone effect), the residuals in the non-spatial model showed a much weaker, albeit significant, spatial autocorrelation based on the Moran's I.

Discussion

Spatial autocorrelation in the rainfall seasonality index accounted for most of the spatial autocorrelation in the malaria seasonality index, as shown by the comparison of the Moran's I of the raw data with the Moran's I of the residuals of a non-spatial model including the rainfall seasonality index. The additional spatial correlation could be due to unobserved variables that change gradually over space, such as those related to soil conditions and altitude (temperature), or factors intrinsic to malaria transmission (as nearby districts may influence other districts) or due to cross-border reporting of cases. However, the analysis excluding those areas that were thought particularly affected by cross-border reporting yielded a similar spatial correlation index. The CAR(γ, τ_s) model showed the lowest DIC. However, there are some doubts as to whether this model really performed this well as the gamma parameter was not significantly different from zero. Nevertheless, since the regression models indicated that the 95% credibility intervals for the coefficient for rainfall did not include zero in the spatial models, there was evidence for a significant correlation between seasonality of rainfall and malaria.

The work most similar to the work presented here is that by Mabaso et al. [52,92] who made use of a seasonal concentration index to summarize seasonality in malaria incidence and entomological inoculation rate (EIR). The seasonal concentration index is based on vector representation (i.e., both magnitude and direction) of the mean monthly values in a given year. Mabaso et al. [92] found that the seasonal concentration index of rainfall was significant in explaining the seasonal concentration index of EIR across Africa. The EIR is more closely related to environmental variables and is a better measure of the risk of inoculation than reported malaria cases, which is confounded by immunological processes. When

71

protective immunity is high as a result of high levels of transmission, as is the case in many African settings, it strongly confounds the relationship between the (all age) incidence time-series and transmission [15]. When immunity plays an important role in disease transmission, it may create and maintain so called "endogenous cycles" in incidence time-series, even when the vectorial capacity (the mosquito vector population's potential to transmit malaria) is at a constant level. Cyclical patterns in incidence time-series may thus partly be caused by immunity dynamics and this may confound the relationship between incidence and extrinsic drivers such as weather [16]. Because of extremely low sporozoite rates in vector mosquitoes in Sri Lanka, it is difficult to measure the EIR which is otherwise a good measure of the risk of inoculation. In the absence of reliable EIR estimates, the malaria-case incidence might be a satisfactory measure of parasite transmission, provided that protective and anti disease immunity is low. Although there are strong similarities between the seasonal concentration index and the seasonality index presented here, there are also important differences. For example, neither differentiate between bimodal systems with (evenly spaced) peaks of similar height and systems without seasonality. However, the seasonality index is continuous, whereas the seasonal concentration index is contained in the zero–one (zero–100 if expressed as percentage) interval. Most importantly, in bimodal systems, the seasonality index allows differentiating between a situation with the first peak being higher than a second (and the reverse situation), whereas the seasonal concentration index does not differentiate.

In this paper, seasonality in temperature was not studied as it was considered of less importance than rainfall seasonality in the Sri Lankan context. The country is situated close to the equator and the temperature therefore fluctuates over a narrow range. Also, a large part of its temporal variability is governed by rainfall. Moreover, for most malarious areas (except in the hilly part of the country situated just south of the centre of the island), the temperature (varying between 22 and 32 degrees Celsius) is well within the range suitable for malaria transmission [5]. A study in Ethiopia found temperature to be generally insignificant in explaining malaria for districts below 1,650 m [105]. However, it merits investigation whether temperature influences malaria seasonality in the hills.

This paper provides evidence that, (even) after correction for spatial autocorrelation in the data, rainfall seasonality is significant in explaining malaria seasonality in space.

This suggests that high cross-correlations between rainfall time-series and malaria time-series found elsewhere [96] are not accidental but that rainfall is a driver of seasonality of malaria cases. Rainfall seasonality could thus in theory be used as a predictor of the seasonality of malaria transmission in the absence of malaria case data or EIR data in areas of low transmission and temperatures which are conducive to malaria transmission year round.

Acknowledgements

Chapter 6 – Temporal correlation between malaria and rainfall in Sri Lanka

*This chapter was published in the Malaria Journal 2008, **7:**77*

Olivier JT Briët [*1,2], Penelope Vounatsou[2], Dissanayake M Gunawardena[3], Gawrie NL Galappaththy[4], Priyanie H Amerasinghe[5]

1 International Water Management Institute, P.O. Box 2075, Colombo, Sri Lanka, Telephone: +94112787404, Fax: +94112786854

2 Swiss Tropical Institute, Socinstrasse 57, P.O. Box CH-4002, Basel, Switzerland

3 US Agency for International Development, P.O. Box 7856, Kampala Uganda

4 Anti Malaria Campaign, Head Office Colombo, Sri Lanka

5 International Water Management Institute Sub Regional Office for South Asia, c/o ICRISAT, Patancheru, AP 502 324, Andhra Pradesh, India

* Corresponding author

Abstract

Background

Rainfall data have potential use for malaria prediction. However, the relationship between rainfall and the number of malaria cases is indirect and complex.

Methods

The statistical relationships between monthly malaria case count data series and monthly mean rainfall series (extracted from interpolated station data) over the period 1972 – 2005 in districts in Sri Lanka was explored in four analyses: cross-correlation; cross-correlation with pre-whitening; inter-annual; and seasonal inter-annual regression.

Results

For most districts, strong positive correlations were found for malaria time series lagging zero to three months behind rainfall, and negative correlations were found for malaria time series lagging four to nine months behind rainfall. However, analysis with prewhitening showed that most of these correlations were spurious. Only for a few districts, weak positive (at lags zero and one) or weak negative (at lags two to six) correlations were found in prewhitened series. Inter-annual analysis showed strong negative correlations between malaria and rainfall for a group of districts in the centre-west of the country. Seasonal inter-annual analysis showed that the effect of rainfall on malaria varied according to the season and geography.

Conclusions

Seasonally varying effects of rainfall on malaria case counts may explain weak overall cross-correlations found in prewhitened series, and should be taken into account in malaria predictive models making use of rainfall as a covariate.

Background

Malaria is a complex disease and its transmission and prevalence is influenced by many factors, amongst which (variability in) climatic conditions are considered to play a major role. With increasing weather variability and ability to forecast weather, there is an interest in developing systems for malaria forecasting that incorporate weather related factors as explanatory variables. Many studies in various parts of the world have linked malaria time series to weather variables such as rainfall, temperature and humidity. For instance, by using polynomial distributed lag models, Teklehaimanot and colleagues [105] found that malaria was associated with rainfall and minimum temperature (with the strength of the association varying with altitude) in Ethiopia. Worrall and colleagues [106] used rainfall and maximum temperature at a lag of four months to successfully fit a biological transmission model to malaria case data in a district in Zimbabwe. Craig and colleagues [107] linked inter-annual differences in malaria to rainfall and temperature in South Africa.

Sri Lanka has a long history of researching the links between rainfall and malaria and many studies observed links between the two [7,8,11,108], but others did not find a strong [109] or an obvious correlation [110]. A study in Sri Lanka incorporating rainfall as a linear or non linear explanatory variable into a (seasonal) auto-regressive integrated moving average (ARIMA) model showed little improvement in malaria prediction over ARIMA models without a rainfall predictor [111].

Weather affects the malaria incidence mostly through its effects on both the mosquito vector (species, population dynamics, gonotrophic cycle and survivorship [112]) and the development of the malaria parasite inside the mosquito vector. In Sri Lanka, the main malaria vector *Anopheles culicifacies* breeds primarily in river bed pools [2] which occur during dry periods, but also in other breeding sites such as seepage areas next to irrigation tanks, hoof prints, and abandoned gem mining pits. The spatial variation in annual precipitation (Figure 18) has been linked to spatial variation in malaria endemicity in Sri Lanka [61] by early malariologists who used the concept of a classification of the country into a wet, intermediate and dry zone [7] based on the amount of rainfall received during the south-west monsoon. The region receiving the most annual precipitation has the least malaria, and endemicity increases with decreasing annual rainfall. The fact that the districts in the extreme south west of the island (Galle and Kalutara) have always been virtually free of malaria is attributed to

the wet climate in which rivers flow year round without pooling. In the south west, only a drought might cause pooling in rivers and hence create conditions suitable for the breeding of *An. culicifacies*. For example, districts with wet and intermediate annual rainfall in this region have repeatedly been affected by malaria epidemics, mostly attributed to droughts due to a failing south-west monsoon (which occurs normally between February and July), while districts towards the north and east with dryer climates (and with a higher malaria endemicity) were less affected [7]. Hence a negative correlation between rainfall and malaria is expected in districts in the wet and intermediate rainfall zones. In contrast, towards the north and east, where the climate is much dryer (particularly during April – September) and rivers often run dry, rainfall creates new puddles, especially following a period of drought. The relation between rainfall and malaria may not only change over space, but also depending on the season. Both rainfall and malaria show a marked seasonality which is bimodal in the south west and increasingly monomodal towards the north and east [113] (Figure 18 and Figure 19). There is a very interesting duality in the relationship of rainfall with malaria. Although malaria prevalence (not shown in this paper, see *e.g.* [61]) decreases with annual rainfall in space, study of the seasonality of both (Figure 18 and Figure 19) reveals that malaria follows rainfall with a few months delay, a low seasonal rainfall peak being followed by a low seasonal malaria peak, and a high seasonal rainfall peak being followed by a high seasonal malaria peak. In the present paper, the relationship between rainfall and malaria incidence in Sri Lanka was investigated allowing for spatial variability in the relationship using (i) cross-correlation analysis, (ii) cross-correlation analysis with prewhitening (prior removal of seasonality and auto-correlation in the series), (iii) inter-annual analysis and (iv) seasonal inter-annual analysis allowing for seasonal variability in the relationship. Only the first of these four approaches has been used to study malaria and rainfall relationships in Sri Lanka previously, and only in limited geographic areas. A better understanding of the relationship between rainfall and malaria may help to improve forecasting of changes in malaria incidence.
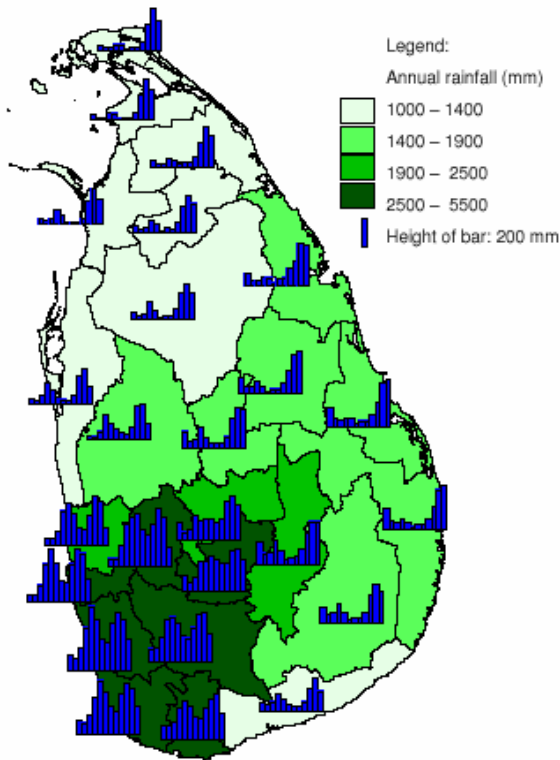
Figure 18 - Geometric mean seasonality and annual geometric mean total of rainfall

Geometric mean monthly rainfall from January (bar on far left) to December (bar on far right), calculated over the period January 1971 to December 2005, and the total geometric mean annual rainfall in districts of Sri Lanka. The height of the bar in the legend represents 200 mm. The classification of the annual rainfall follows the common delineation into a wet zone (>2500 mm per annum), intermediate zone, and a dry (<1900 mm per annum) zone, with a "very dry" category for rainfall <1400 mm per annum.
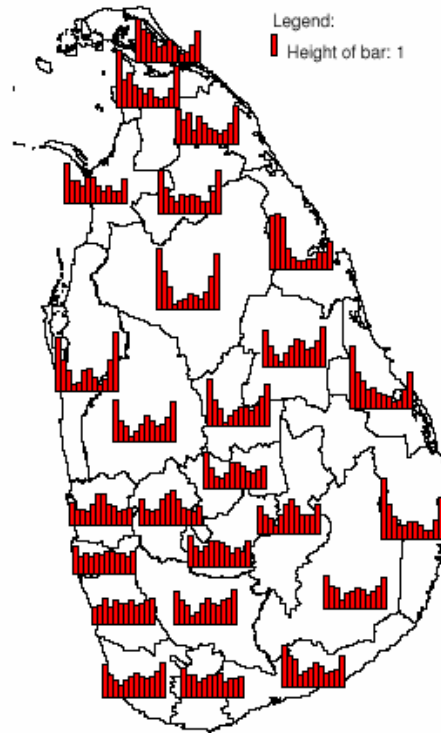
Figure 19 - Geometric mean seasonality of detrended malaria cases

Geometric mean monthly number of malaria cases from January (bar on far left) to December (bar on far right), calculated over the period January 1972 to December 2005, after detrending, in districts of Sri Lanka. The height of the bar in the legend represents 1, (because of detrending, no unit is given).

Methods

Data

Since January 1972, the Sri Lankan national malaria control programme, the Anti
Malaria Campaign (AMC), has collected monthly confirmed malaria case data from
health facilities aggregated by medical officer of health (MOH) areas (which
represent sub district health administrative divisions). This data up to December 2005
was cleaned and aggregated to district resolution. For each district, for each month,
the mean rainfall was extracted from monthly rainfall surfaces for the period January
1971 – December 2005. Both rainfall and malaria datasets are described in detail
elsewhere [111].
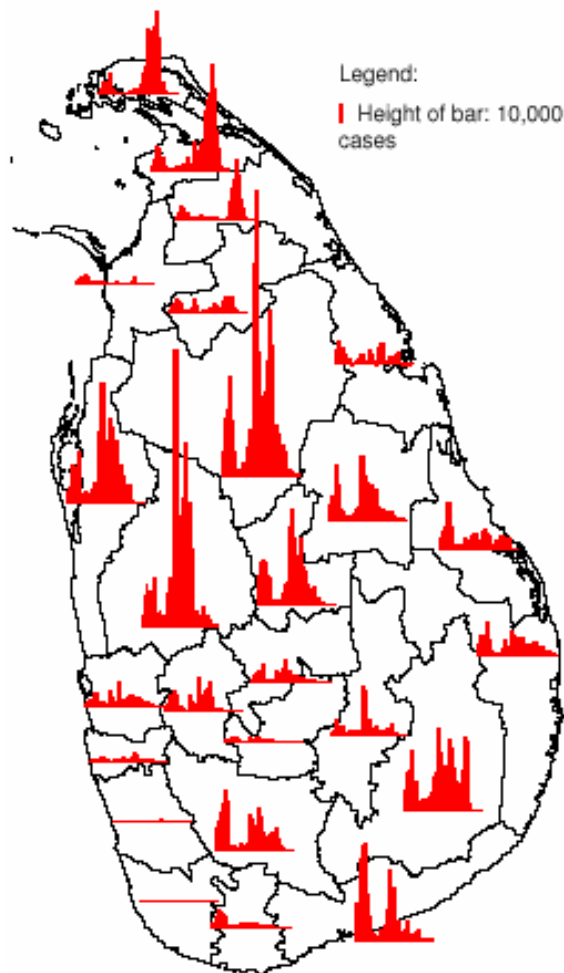
Figure 20 - Annual malaria cases

Annual number of malaria positive cases from 1972 (bar on far right) to 2006 (bar on for left) in
districts in Sri Lanka. The height of the bar in the legend represents 10,000 cases.

Statistical analysis

The relationship between rainfall and malaria incidence was investigated using (i) cross-correlation analysis, (ii) cross-correlation analysis with prewhitening, (iii) inter-annual analysis and (iv) seasonal inter-annual analysis allowing for temporal variability in the effect.

Cross-correlation analysis

Cross-correlations between detrended monthly malaria case count time series and monthly total rainfall [109] were analysed to detect the time lag(s) of rainfall preceding malaria at which the series show strongest correlation.

Malaria time series showed strong long term fluctuations for most districts in Sri Lanka (Figure 20). However, in rainfall time series these long term fluctuations were absent. Therefore, it was expected that rainfall could not explain the long term fluctuations in malaria, which were probably related to other factors, such as malaria control and population changes. The long term fluctuations masked the correlation between malaria and rainfall and since no information on the underlying factors was available in the data, the long term fluctuations needed to be removed prior to calculating cross-correlations. It was assumed that untransformed monthly malaria case count data follow a seasonal model of the form:

$$y_t = m_t S_t \varepsilon_t$$

where y_t is the malaria case count in month t; m_t is the mean level in month t; S_t is the seasonal effect in month t; and ε_t is the random error. A logarithmic transformation $y'_t = \log(y_t + 1)$ was employed to transform the malaria case counts to normality and turn equation (1) into a (linear) additive model [90] of the form:

$$y'_t = m'_t + S'_t + \varepsilon'_t \text{ where } m'_t \approx \log(m_t), \ S'_t \approx \log(S_t) \text{ and } \varepsilon'_t \approx \log(\varepsilon_t).$$

As an example, Figure 21 shows the logarithmically transformed series for Gampaha district. The long term fluctuations m'_t in the logarithmically transformed monthly district malaria case count series were calculated using a 13-point centred smoothing filter with the months at the extremes given half weight:

$$m_t = \tfrac{1}{12}\left(\tfrac{1}{2} y'_{t-6} + y'_{t-5} + \ldots + y'_t + \ldots + y'_{t+5} + \tfrac{1}{2} y'_{t+6} \right).$$

Smoothing was performed using the function "decompose" of the package "stats" in the software R [56]. From the detrended series $\zeta_t = y'_t - m'_t$ (Figure 22) implicitly long term trends caused by population growth were removed. Cross-correlation analysis was applied between the detrended log transformed malaria case time series and untransformed rainfall time series x_t. The cross-correlation was estimated for malaria with a lag l of zero to twelve months behind rainfall as

$$r_l = \sum_{t=1}^{N} (x_t - \bar{x})(\zeta_{t+l} - \bar{\zeta}) / N s_x s_\zeta$$

where s_x, s_ζ are the sample standard deviations of observations on x_t and ζ_t, respectively. The analysis was repeated with logarithmically transformed rainfall time series $x'_t = \log(x_t + 1)$.

Even though the above approach may find strong correlations, these may not be very useful for malaria prediction if aberrations from the long term seasonal mean of rainfall are weakly linked to aberrations from the long term seasonal mean of the malaria case series. In addition, the standard cross-correlation assumes observations are independent, whereas in reality the malaria data are temporally correlated.
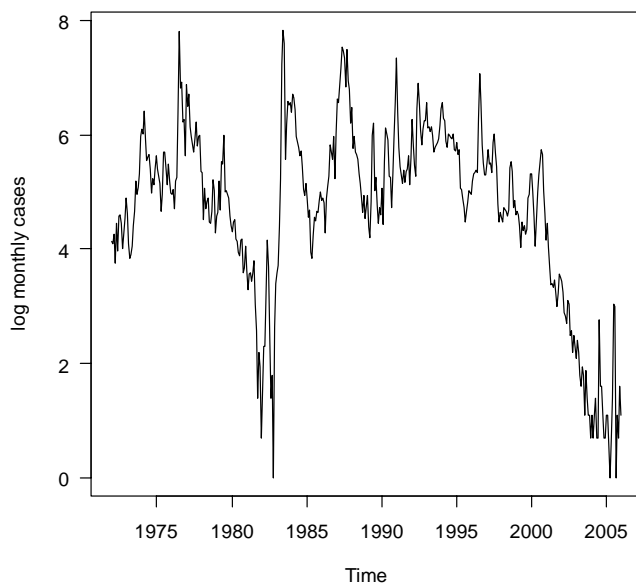
Figure 21 - Logarithmically transformed monthly malaria case counts for Gampaha District

Logarithmically transformed monthly malaria case counts (after adding the value 1 to all data) for Gampaha District.
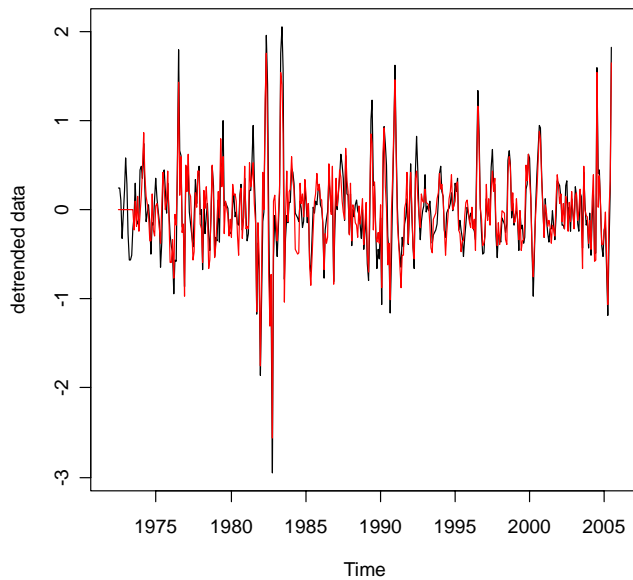
Figure 22 - Detrended (prewhitened) logarithmically transformed monthly malaria case counts for Gampaha District

Detrended logarithmically transformed monthly malaria case counts (black line) and prewhitened detrended logarithmically transformed monthly malaria case counts (red line) for Gampaha District.

Cross-correlation analysis with prewhitening

Cross-correlation with the seasonality and autocorrelation removed by simple prewhitening allows for detection of the time lag(s) of rainfall preceding malaria, at which divergences from the long term seasonal pattern in rainfall time series shows strongest correlation with such divergences in detrended malaria case count time series, while minimizing effects of spurious correlations caused by autocorrelation in the time series. This method bears some similarity to anomaly analysis, where the cross-correlation of aberrations from the long term seasonal mean of the explanatory variables is weakly correlated with aberrations from the long term seasonal mean of the response variable. It should be noted that in this analysis, the cross correlation is calculated as the average over all (calendar) months. Possible varying correlation depending on the season is not accounted for, *i.e.* if rainfall has a strong positive effect on malaria in some months, and a strong negative in others, the average detected cross correlation could be weak. The effect of prewhitening is to reduce unassociated autocorrelation and/or trends within time series prior to computation of

their cross correlation function (It is well established that autocorrelation within time series can produce spurious cross correlations [90]). Simple prewhitening is used when there is a clear unidirectional influence such as between rainfall and malaria. First, an auto-regressive model is fit to the explanatory variable. The prewhitened explanatory variable consists of the residuals of the fitted model. The prewhitened outcome variable consists of the residuals of the same model (with the same parameters) applied to the outcome variable. With the inclusion of seasonality in the autoregressive model, the prewhitening procedure removes seasonality (and autocorrelation) from the explanatory variable time series, and the same amount of seasonality (and autocorrelation) from the outcome variable time series. It is thus possible that additional seasonality (and autocorrelation) remains in the prewhitened outcome variable time series.

For each district, multiplicative seasonal auto-regressive integrated moving average (SARIMA) models [100] with all possible combinations of parameters p, q, P, Q $\in$ {0, 1, 2} and with d, D $\in$ {0, 1}, were evaluated using the Akaike's information criterion (AIC) on untransformed and logarithmically transformed monthly rainfall totals in the period from January 1971 to December 2005.

The selected SARIMA model was then used to prewhiten both the rainfall time series and detrended (smoothed) logarithmically transformed malaria case count time series ζ_t. The residuals of both series were used as input for the cross-correlation analysis. The functions "arima" and "ccf" from the R package "stats" were used.

The cross-correlation analyses above have the drawback of masking inter-annual effects of rainfall on malaria time series because of the removal of the strong long term trend fluctuations.

Inter-annual analysis

In "Inter-annual analysis", the series of differenced logarithmically transformed annual malaria cases was studied to determine if it was correlated to differenced logarithmically transformed total annual rainfall. Unlike the first two approaches, it could not account for the within year effects, but inter-annual effects were not masked.
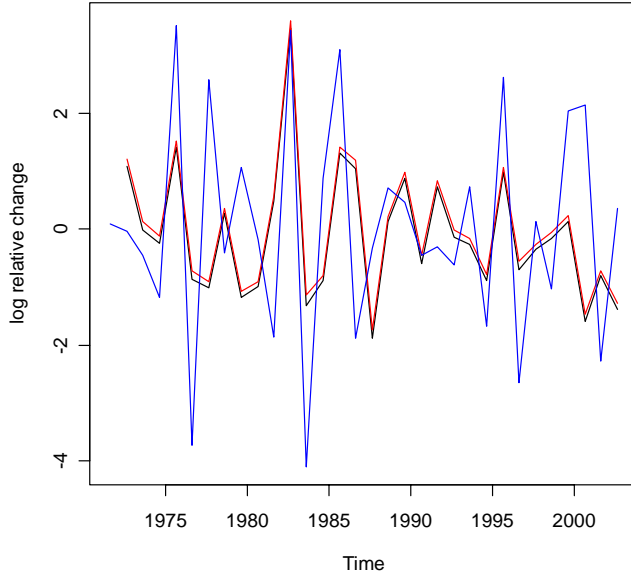
Figure 23 - Differenced logarithmically transformed annual malaria case counts and rainfall for Gampaha District

Differenced logarithmically transformed annual (the twelve month period starting in April) malaria case counts (black line), malaria case counts with first order auto correlation removed (red line), and the differenced logarithmically transformed annual rainfall with a three month lag shift (the twelve month period starting in January), corrected for autocorrelation in malaria and multiplied by -10 (blue line) for Gampaha District.

The difference $\Omega_{t,k} = \log(Y_{t,k}) - \log(Y_{t-1,k}) = \log(Y_{t,k}/Y_{t-1,k})$ reflects the relative change in case numbers between consecutive years [107], where $Y_{t,k}$ is the annual malaria case total for year t, and the start month k of the twelve-month period was either April ($k = 4$) or September ($k = 9$) because seasonally, malaria was lowest in either April or September, depending on the district [61]. Similarly, the relative change in rainfall over 12 month periods preceding the malaria periods with a lag l of one to three months was represented by

$\Xi_{t,l,k} = \log(X_{t,k,l}) - \log(X_{t-1,k,l}) = \log(X_{t,k,l}/X_{t-1,k,l})$. Malaria was regressed against rainfall in a first order auto-correlated (AR1) model:

$\Omega_{t,k} = \phi_k \Omega_{t-1,k} + \beta_{l,k}(\Xi_{t,l,k} - \phi_k \Xi_{t-1,l,k}) + \varepsilon_{t,k}$. The Pearson correlation coefficient between $\Omega_{t,k} - \phi_k \Omega_{t-1,k}$ and $\Xi_{t,k,l} - \phi_k \Xi_{t-1,k,l}$ was then calculated. Figure 23 and Figure

24 provide an illustration for Gampaha District. The robustness of all significant ($p \leq 0.1$) correlations detected was tested as follows: For each observation it was calculated whether it was influential in terms of dfbeta, dffits, covariance ratios, Cook's distances and the diagonal elements of the hat matrix. Observations which were influential with respect to any of these measures were omitted (one at a time) from the data and the correlation coefficient was recalculated. The weakest correlation among these was recorded.
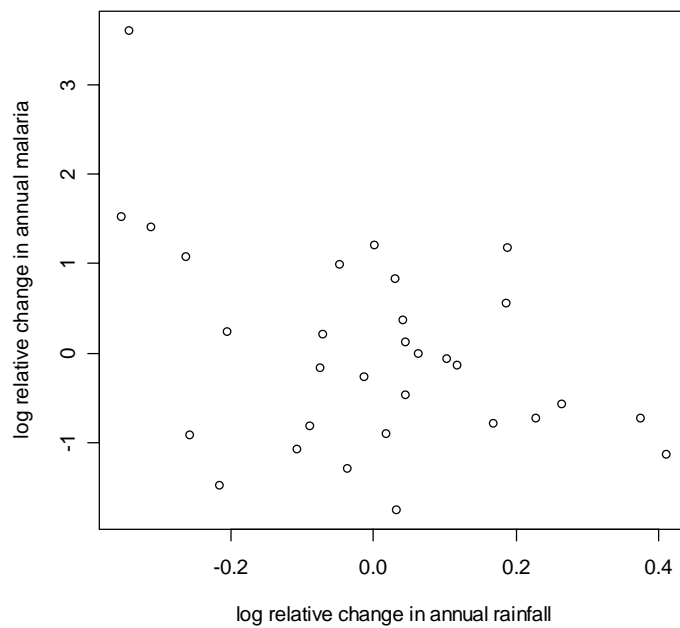
Figure 24 - Scatter plot of differenced logarithmically transformed annual malaria case counts and rainfall for Gampaha District

Scatter plot of differenced logarithmically transformed annual (the twelve month period starting in April) malaria case counts with first order auto correlation removed against differenced logarithmically transformed annual (the twelve month period starting in January) rainfall corrected for first order auto correlation in malaria for Gampaha District.

Seasonal inter-annual analysis

The effect of rainfall on malaria may depend on the season; therefore, it was of interest to assess the inter-annual relationship between malaria and rainfall for each calendar month in the year. The inter-annual analysis above was modified by replacing $\Omega_{t,k}$ with $\omega_{t,k}$, and $\Xi_{t,k,l}$ with $\xi_{t,k,l}$. Here, $\omega_{t,k}$ represents the average

logarithmically transformed malaria count over three months (*e.g.* January - March) differenced with the average logarithmically transformed malaria in the previous twelve months: $\omega_{t,k} = \sum_{k-1}^{k+1} \log(y_{t,k}+1)\Big/3 - \sum_{k-13}^{k-2} \log(y_{t,k}+1)\Big/12$ with $y_{t,k}$ the malaria count in month k (varied between January and December) and in year t. Similarly, $\xi_{t,k,l} = \sum_{k-1-l}^{k+1-l} \log(x_{t,k,l}+1)\Big/3 - \sum_{k-13-l}^{k-2-l} \log(x_{t,k,l}+1)\Big/12$. The seasonally varying correlation coefficients between rainfall and malaria $r_{k,l}$ were transformed into $z_{k,l}$ values using the Fisher transformation $z_{k,l} = 0.5\log\left(\dfrac{1+r_{k,l}}{1-r_{k,l}}\right)$ and correlated to a three month centred moving average of logarithmically transformed geometric mean seasonal rainfall (similar as depicted in Figure 18, but logarithmically transformed) and its derivative (expressing the change in seasonal rainfall per month).
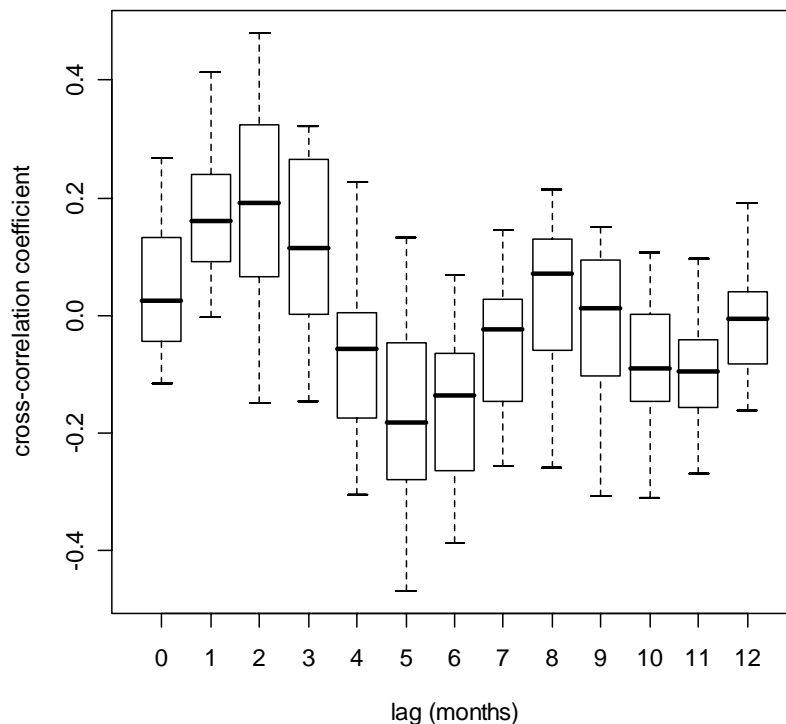
Figure 25 - Cross-correlation box plot

Box plot of Pearson product-moment correlation coefficients of time series of logarithmically transformed monthly rainfall and (detrended) monthly logarithmically transformed malaria case time series at several lags for districts in Sri Lanka, grouped by lag distance.

Results

Cross-correlation analysis

For all districts, a local maximum cross-correlation between malaria and
untransformed rainfall or logarithmically transformed rainfall was found when rainfall
was preceding malaria by zero to three months, depending on the district. For 13 out
of 25 districts, logarithmic transformation of rainfall improved the cross-correlation
(Figure 25), and for some districts the logarithmic transformation of rainfall caused
the lag of the local peak correlation to shift by a month. For most districts, the
optimum was found at a lag of two months. Neighbouring districts showed similar
cross-correlation coefficients at similar lags (Figure 26). The peak correlation
coefficient was as high as 0.5 for some districts (*e.g.* Anuradhapura and Puttalam), but
very low and not significant for others (*e.g.* Badulla). A local minimum cross-
correlation between malaria and untransformed rainfall or logarithmically transformed
rainfall was found when rainfall was preceding malaria by four to ten months,
depending on the district. For most districts, a second local maximum was found
when rainfall was preceding malaria by seven to nine months.
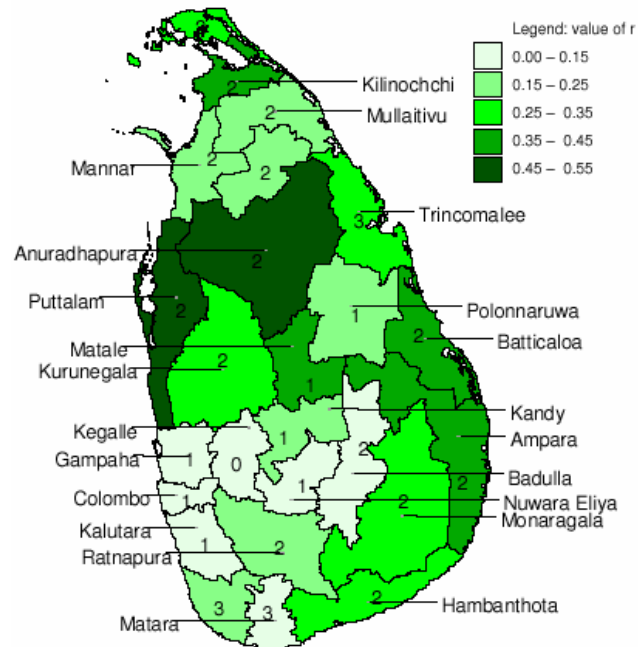
Figure 26 - Mapped maximum cross-correlation coefficients

Mapped maximum cross-correlation coefficients for logarithmically transformed monthly rainfall
preceding (detrended) logarithmically transformed monthly malaria case time series with zero to
twelve months for districts in Sri Lanka. Numbers indicate the lag (in months) for which the maximum
occurred.
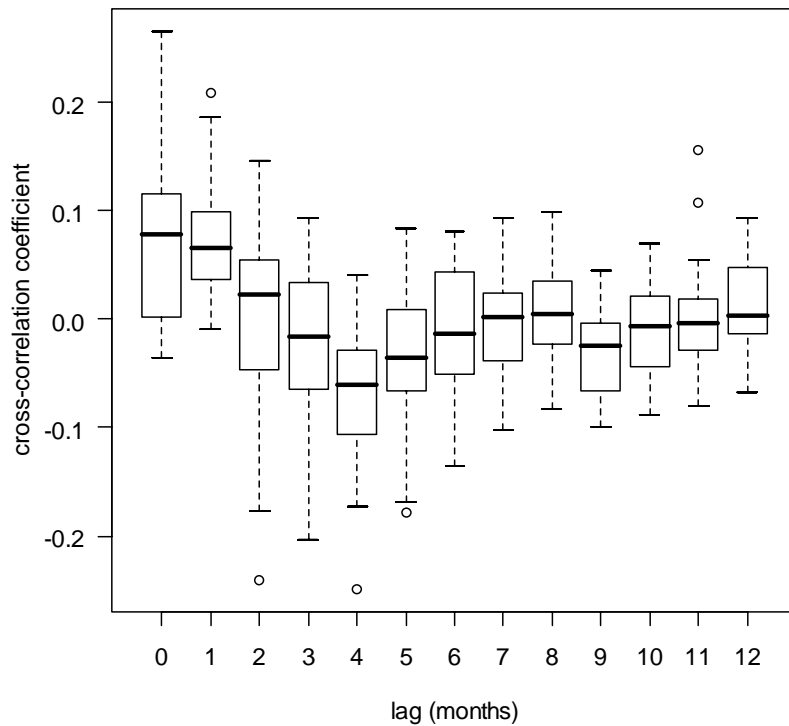
Figure 27 - Cross-correlation box plot after prewhitening (rainfall log-transformed)

Box plot of Pearson product-moment correlation coefficients of prewhitened series of logarithmically transformed monthly rainfall and (detrended) monthly malaria case time series at several lags for districts in Sri Lanka, grouped by lag distance.

Cross-correlation analysis with prewhitening

For prewhitening, the SARIMA models applied to the (logarithmically transformed) rainfall data showed that for all districts, the model with the lowest AIC had a seasonal component (P=0, D=1, Q=1), and results were very similar among all non seasonal components (p, d, q) tested, except for the components (p=0, d=1, q=0), (p=1, d=1, q=0), and (p=2, d=1, q=0), which gave worse results. The model SARIMA(p=1, d=0, q=0, P=0, D=1, Q=1) was selected. Figure 22 shows the effect of prewhitening on the malaria time series for Gampaha district.

With prewhitened time series, the cross-correlograms looked entirely different (Figure 27) from the cross-correlograms without prewhitening (Figure 25). Correlations were generally weaker with prewhitening than without. Like with the analysis without prewhitening, neighbouring districts showed similar cross-correlation coefficients at

similar lags (Figure 28 and Figure 29). For 18 out of 25 districts, logarithmic transformation of rainfall improved the cross-correlation for the first local maximum, and for 12 out of 25 districts, it improved for the local minimum following the first maximum, as compared to untransformed rainfall. The prewhitened series showed strongest positive correlations at lags of zero and one month, and only for five out of 25 districts (Puttalam, Kurunegala, Matale, Kegalle and Moneragala, all neighbouring districts except Moneragala) the correlation coefficient was over 0.15. Strongest negative associations were found at lags of two to five months, and only for six out of 25 districts (Gampaha, Kegalle, Kurunegala, Matale, Nuwara Eliya and Ratnapura, all adjoining districts) the correlation coefficient was below -0.15.
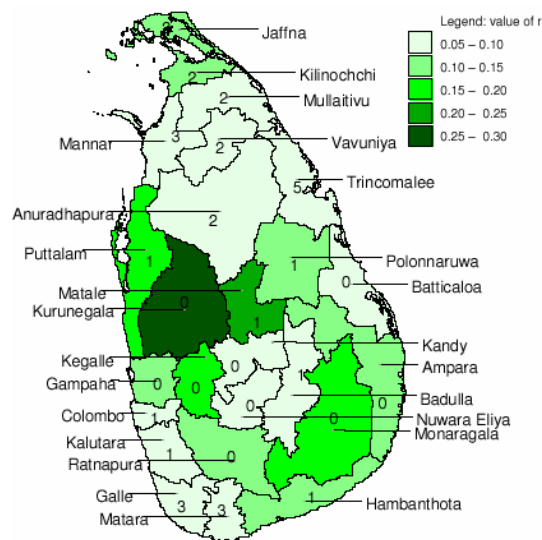
Figure 28 - Mapped maximum cross-correlation coefficients after prewhitening

Mapped maximum cross-correlation coefficients for logarithmically transformed rainfall preceding (detrended) logarithmically transformed malaria case time series with zero to twelve months for districts in Sri Lanka, after prewhitening. Numbers indicate the lag (in months) for which the maximum occurred.
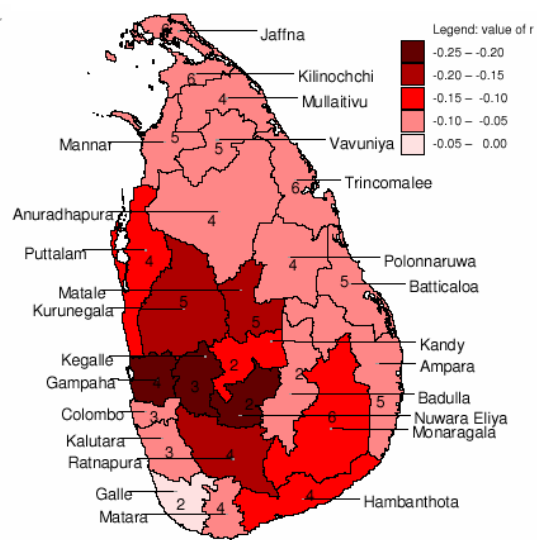
Figure 29 - Mapped minimum cross-correlation coefficients after prewhitening

Mapped minimum cross-correlation coefficients for logarithmically transformed monthly rainfall preceding (detrended) logarithmically transformed monthly malaria case time series with zero to twelve months for districts in Sri Lanka, after prewhitening. Numbers indicate the lag (in months) for which the maximum occurred.

Table 7 - Maximum and minimum Pearson product-moment cross-correlation coefficients, starting month and lag (number of months that malaria case time series are lagged behind) for which the maximum or minimum occurred, and significance of the regression coefficient for logarithmically transformed rainfall and differenced logarithmically transformed annual malaria case time series (n = 32), corrected for first order auto regressive correlation.

District	Minimum		Maximum	
	cc	start month (lag)	cc	start month (lag)
Ampara	-0.01	4(3)	0.23	9(1)
Anuradhapura	0.08	4(1)	0.27	9(3)
Badulla	-0.09	4(1)	0.30	9(1)
Batticaloa	-0.06	9(3)	0.23	4(1)
Colombo	-0.05	4(3)	0.25	9(3)
Galle	-0.27	4(1)	0.27	9(3)
Gampaha	-0.41*	4(3)	0.03	9(3)
Hambantota	-0.24	4(3)	0.08	9(3)
Jaffna	-0.16	4(3)	0.13	9(1)
Kalutara	-0.10	4(3)	0.25	9(2)
Kandy	-0.40*	4(3)	-0.09	9(3)
Kegalle	-0.55**	4(3)	-0.13	9(3)
Kilinochchi	-0.03	4(2)	0.12	9(1)
Kurunegala	-0.32'	4(3)	0.13	9(1)
Mannar	-0.08	4(3)	0.06	4(1)
Matale	-0.18	4(1)	0.10	9(1)
Matara	-0.11	4(2)	0.13	9(2)
Moneragala	-0.17	4(3)	0.23	9(1)
Mullaitivu	-0.20	9(3)	0.06	4(3)
Nuwara Eliya	-0.36*	9(2)	-0.23	9(3)
Polonnaruwa	-0.22	4(3)	0.17	9(1)
Puttalam	-0.14	4(3)	0.18	9(1)
Ratnapura	-0.52**	4(3)	-0.03	9(1)
Trincomalee	-0.30'	9(3)	-0.07	4(3)
Vavuniya	-0.35'	(4)2	-0.26	(4)1

Legend: cc=Pearson product moment correlation coefficient, significance of regression coefficient different from zero: ' = P<0.10,* = P<0.05, ** = P<0.01

Inter-annual analysis

None of the districts showed significant ($p \leq 0.1$, n = 32) positive correlation coefficients (Table 7), and eight districts showed significant negative correlation coefficients (Figure 30). After omitting the influential observation that contributed most to the correlation, for four adjoining districts the correlation coefficients were still significant ($p \leq 0.1$, n = 31). These districts were Kurunegala (r = -0.32, p = 0.08), Kegalle (r = -0.48, p = 0.007), Kandy (r = -0.34, p = 0.063), and Ratnapura (r = -0.52, p = 0.003).
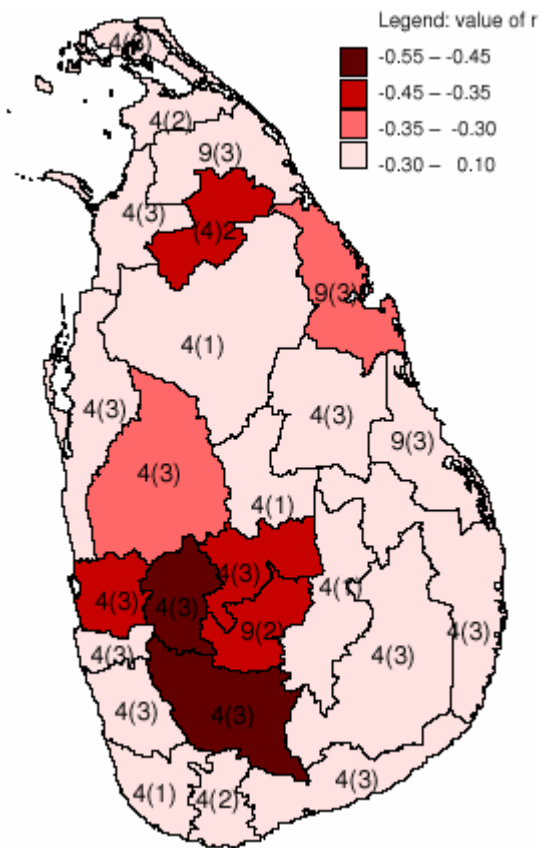
Figure 30 - Mapped minimum inter-annual cross-correlation coefficients

Mapped minimum cross-correlation coefficients for logarithmically transformed annual rainfall preceding differenced logarithmically transformed annual malaria case time series (with first order auto correlation removed (see methods)), with one to three months for districts in Sri Lanka. Numbers indicate the starting month of the year (4 = April, 9 = November) and between brackets the lag (in months) for which the minimum occurred.

Seasonal inter-annual analysis

In a given district, rainfall over a particular three month period (*e.g.* January –
March), relative to rainfall in the preceding twelve month period, had in general a
similar effect on the malaria count over three months, relative to the malaria in the
preceding twelve month period, for malaria following rainfall with a time lag of one
(*e.g.* malaria in February – April) to three (*e.g.* malaria in May – June) months,
although cross correlations were stronger positive at a lag of one month and stronger
negative at a lag if three months. This is illustrated for the district of Gampaha in
Figure 31. The cross-correlation coefficient for rainfall preceding malaria with a lag
of two months is presented in Figure 32. This figure shows strong negative
correlation coefficients for districts in the centre west of the country for rainfall
during February – June. After omitting the influential observation that contributed
most to the correlation, the troughs were significant ($p \leq 0.1$, n $= 31$) for the districts
Gampaha, Kegalle and Nuwara Eliya. For Puttalam, Gampaha and Kegalle,
significant peaks were observed at the end of the year. In the north and east, some
districts showed positive correlation during the middle of the year (significant peaks
for the districts Jaffna, Kilinochchi, Batticaloa, Ampara and Moneragala). Jaffna
District showed significant positive peaks at the end of the year, whereas for close by
Mullaitivu District the relationship was negative. The Fisher transformed seasonal
inter-annual correlation coefficients at a lag of two months were significantly
($p \leq 0.1$, n $= 12$) negatively correlated to seasonal rainfall in some districts in the east
(Mullaitivu, Mannar and Polonnaruwa, Figure 33 and Figure 34), whereas there was a
positive correlation for the districts Anuradhapura, Kandy, Nuwara Eliya and
Kalutara. A more smooth picture was obtained by correlating the derivative of
seasonal rainfall to the Fisher transformed seasonal inter-annual correlation
coefficients. Districts in the centre-west (Gampaha, Kegalle and Colombo) and in the
north (Jaffna) showed significant negative correlations (Figure 35 and Figure 36).
Districts in the east showed positive correlation (significant for Mullaitivu, Batticaloa
and Ampara). Galle also showed a significant positive correlation, but these results
should be interpreted with the knowledge that the few infections recorded there are
presumed to have been acquired elsewhere.
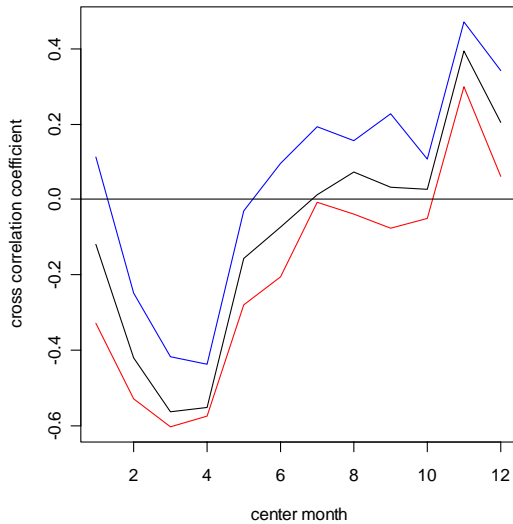
Figure 31 - Cross-correlation coefficients for each rainfall month with malaria lagging one to three months behind for Gampaha District.

Cross-correlation coefficients for logarithmically transformed three-monthly rainfall (differenced with the logarithmically transformed rainfall in the preceding twelve months) with logarithmically transformed three-monthly number of malaria cases (differenced with the logarithmically transformed number of malaria in the preceding twelve months), after removing first order auto correlation (see methods)), with the malaria series lagging one (blue line), two (black line) and three (red line) months behind the rainfall series, for the districts of Gampaha in Sri Lanka. The time scale on the horizontal axis reflects the centre month for three rainfall months.
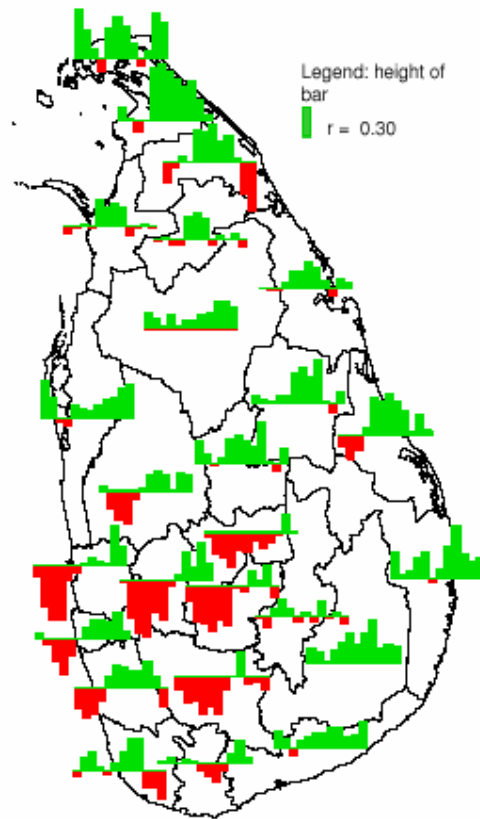
Figure 32 - Mapped seasonal cross-correlation coefficients for malaria lagging two months behind rainfall.

Mapped cross-correlation coefficients for logarithmically transformed three-monthly rainfall (differenced with the logarithmically transformed rainfall in the preceding twelve months) with logarithmically transformed three-monthly number of malaria cases (differenced with the logarithmically transformed number of malaria in the preceding twelve months), after removing first order auto correlation (see methods)), with the malaria series lagging two months behind the rainfall series, for districts in Sri Lanka. The bar on the far left represents January as the centre month of a three months rainfall period; the bar on the far right represents December. Red bars represent negative correlation, green bars represent positive correlation.
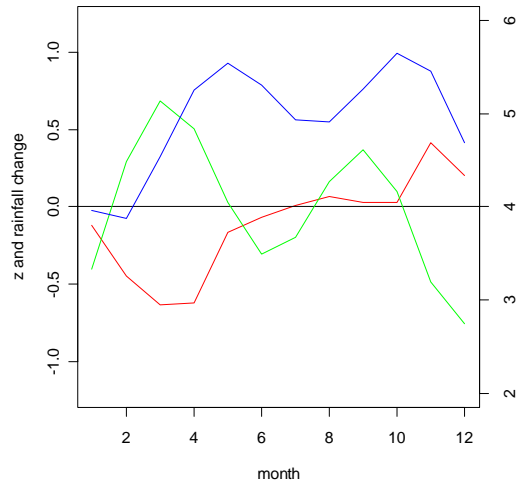
Figure 33 - Correlation coefficients and rainfall for Gampaha District

Three month centred moving average of logarithmically transformed geometric mean monthly rainfall (in mm per month, calculated over the period January 1971 to December 2005) (blue line on right vertical axis), its derivative representing logarithmically transformed rainfall change per month (green line on left vertical axis) and the Fisher transformed correlation coefficient (red line on left vertical axis) between malaria and rainfall at a lag of two months, found in seasonal inter-annual analysis (see methods) for Gampaha District.
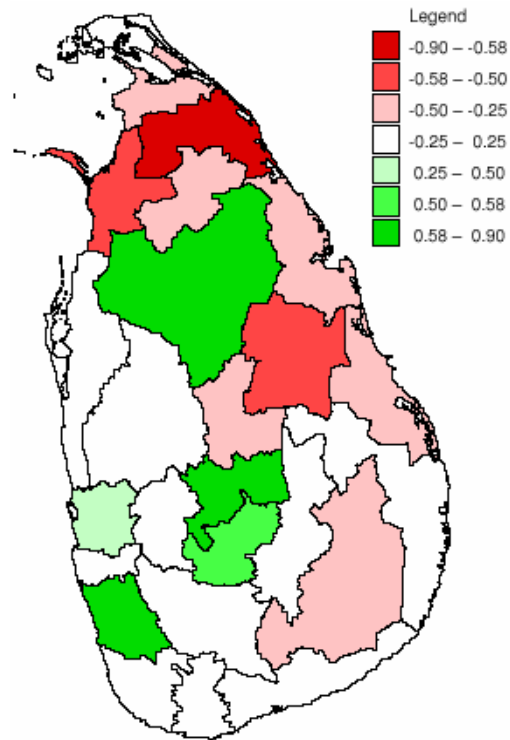
Figure 34 - Correlation between correlation coefficients and rainfall for districts in Sri Lanka

Mapped correlation coefficient between the Fisher transformed correlation coefficient between malaria and rainfall found in seasonal inter-annual analysis at a lag of two months, (see methods) and a three month centred moving average of logarithmically transformed geometric mean monthly rainfall (in mm per month, calculated over the period January 1971 to December 2005) for districts in Sri Lanka

94
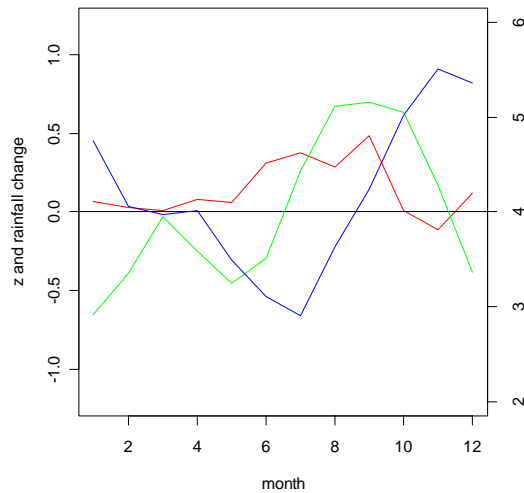
Figure 35 - Correlation coefficients and rainfall for Polonnaruwa District

Three month centred moving average of logarithmically transformed geometric mean monthly rainfall (originally in mm per month, calculated over the period January 1971 to December 2005) (blue line on right vertical axis), its derivative representing change in logarithmically transformed rainfall per month (green line on left vertical axis) and the Fisher transformed correlation coefficient (red line on left vertical axis) between malaria and rainfall at a lag of two months, found in seasonal inter-annual analysis (see methods) for Polonnaruwa District.
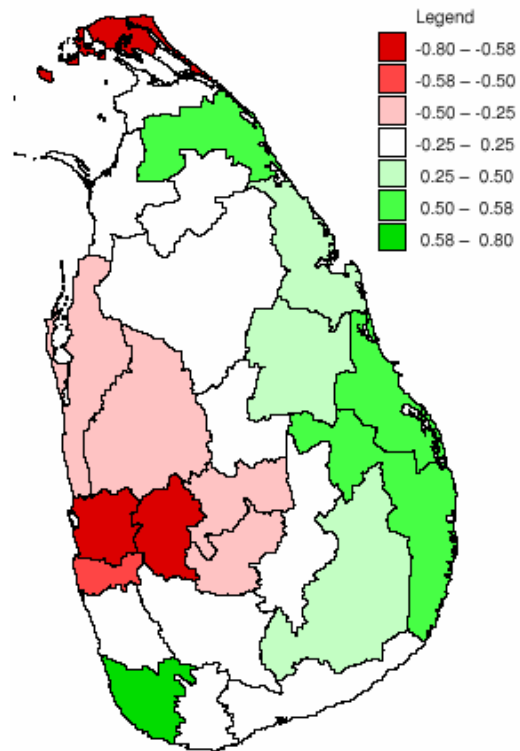
Figure 36 - Correlation between correlation coefficients and change in rainfall for districts in Sri Lanka

Mapped correlation coefficient between the Fisher transformed correlation coefficient between malaria and rainfall found in seasonal inter-annual analysis at a lag of two months, (see methods) and monthly change in a three month centred moving average of logarithmically transformed geometric mean monthly rainfall (originally in mm per month, calculated over the period January 1971 to December 2005) for districts in Sri Lanka.

Discussion

Cross-correlation analysis

In some districts in Sri Lanka, malaria case time series and rainfall showed high cross-correlations at short lags as well as at longer lags. While a causal relationship is biologically plausible at a lag of two to four months, it is increasingly less so at longer lag times. Amerasinghe and colleagues [3] found a lag period of 1.5 months between a peak in abundance of *Anopheles culicifacies* immature forms and a peak in malaria cases, in a village in Anuradhapura District. An additional time lag between rainfall and its effect on breeding conditions, depending on conditions such as soil moisture content, has to be included for the calculation of the rainfall – malaria time lag. For most districts, a positive cross correlation was observed between malaria and rainfall at a lag of two months, confirming the visual impression obtained by studying Figure 18 and Figure 19. It appears, however, that a large part of the detected cross-correlation is due to auto-correlation and concurring cyclical trends, as the cross-correlation analysis with the prewhitened series (discussed below) showed much smaller cross-correlations at short lags and absence of cross correlation at longer lags.

Cross-correlation analysis with prewhitening

For a few districts, (weak) positive cross-correlations were found in prewhitened series with no lag (Kegalle, Kurunegala and Moneragala) and at a lag of one month (Matale and Puttalam). With a lag of one month, short term prediction with a one month horizon would be possible. However, a one month lag seems the absolute minimum for the biological pathway from creating suitable breeding conditions to mosquito development, parasite development in the mosquito, the onset of disease symptoms, and eventually the taking of a blood sample. Nevertheless, in a study in China, log transformed malaria and rainfall showed a maximum (positive) effect for malaria lagging one month behind rainfall, when entered into a regression model together with minimum temperature and fixed quarterly effects for seasonality [114]. Cross-correlations of rainfall contemporary with malaria (at a lag of zero months) are of no value for malaria prediction systems because the total monthly rainfall for the future month needs to be known, unless rainfall can be predicted with high certainty. The (strongest) negative cross-correlations, albeit weak, found in the six adjoining districts at lags of two to five months in the centre-west, are in line with other studies

that showed that this region (except for Nuwara Eliya district which is situated at high altitude) is particularly prone to epidemics when monsoon rains fail [7]. It is difficult to find explanations for the differences in lag time at which the maximum or minimum cross-correlation occurs among (often neighbouring) districts. Factors that could contribute to these differences are saturation levels and water retention of top soils, factors related to differences in malaria endemicity, and differences in temperature (mainly caused by differences in altitude). However, given the generally weak cross-correlations, a large part of the inter-district variation in time-lag of maximum or minimum cross-correlation could have been caused by stochastic noise.

After prewhitening, the cross-correlations found were very weak. Only if rainfall can explain that part of the variation in a malaria time series that cannot be explained by autocorrelation and repetition of seasonal patterns, a rainfall covariate could contribute to a malaria forecasting system. It was only for two out of the six districts (Gampaha and Ratnapura) with strongest negative correlation (situated in the centre west) that Briët and colleagues [111] found some contribution of rainfall to malaria prediction in seasonal ARIMA models at a lag of two months, and they found no improvement for Matale and Puttalam districts at a lag of one month.

Inter-annual analysis

Some studies in neighbouring India [115,116], with comparable total annual rainfall and strong seasonality in rainfall, have tried but failed to find a significant correlation between annual rainfall and malaria. These studies did not consider differencing or detrending the data. A study in Ingwavuma and Ubombo districts in KwaZulu-Natal province in South Africa, with less annual rainfall than Sri Lanka, also failed to find such correlation between annual malaria and rainfall time series, but it did find significant positive correlations between the difference of successive twelve-monthly (July to June, corresponding to the local malaria season) logarithmically transformed malaria case totals and summer (November – March) rainfall (and temperature) [107], while the long term trends were attributed to non-climatic factors [117]. Likewise, a study in Botswana, also with less annual rainfall than Sri Lanka, found a positive correlation between (detrended) annual malaria anomalies and December – February rainfall [118]. In the present study, strong negative correlations were found between differenced annual malaria and rainfall for a contiguous group of districts in the centre-west (with high annual rainfall), and these results were somewhat in line with

the results found in the cross-correlation analysis with prewhitening. This area in particular has been repeatedly affected by malaria epidemics during droughts in the pre-malaria control era [7], and apparently malaria control has not changed this dynamic. Although initially significant negative correlations were detected for the drier districts Vavuniya and Trincomalee, the correlations in these districts were not very robust to influential observations. The data quality in the north-east has been compromised by the armed conflict in the region, and for some districts (particularly Vavuniya) some missing data were imputed. The strong (negative) inter-annual correlations found for the districts in the centre-west provides hope for the development of long term malaria forecasting systems involving long term weather forecasts, provided these systems have sufficient capabilities to predict rainfall anomalies up to a year in advance, which is currently not feasible. It is tempting to attribute the inverse direction of the relationship between rainfall and malaria found in this analysis as compared to the direction found in Southern Africa to the difference in annual rainfall, although other important differences exist, notably in malaria vector species.

Seasonal inter-annual analysis

The results of the seasonal inter-annual analysis supported the theory that rainfall varies in its effect on malaria transmission depending on the season. These effects may cancel out when averaged over the entire calendar year (inherent to the first three approaches studied), and therefore, it seems that malaria forecasting systems incorporating rainfall need to take this seasonally varying effect into account. Note, however, that Briët and colleagues [111] found limited improvement in malaria prediction with a seasonally varying rainfall effect for only three districts.

There was a marked difference in the season-varying effect of rainfall on malaria between the south-western quadrant of the country and the rest of the country. In the south west, the effect was strongly negative during February – June, whereas in the other quadrants, often a positive effect was found during April – September. In most districts (except in the north-eastern quadrant), also a (weak) positive effect was found in December or January.

Similar to the explanation of the spatial variation in malaria endemicity by spatial variation in annual rainfall, the spatial variation in the (seasonally varying) effect of

rainfall on malaria may be explained by spatial variation in (seasonal) rainfall. In the south west, rainfall is normally lowest between November – April, in contrast with the rest of the country, where the April – September trough is (much) deeper in the rainfall climatology (Figure 18). The classic biological explanation for the epidemics in the centre-west of the country is that failure of the south-west monsoon (that normally occurs between February – July, affecting mostly the wet and intermediate zones in the west) will cause the already low rivers (relative to the rest of the year) to stagnate and create breeding sites for *An. culicifacies*. Thus the strong negative correlation found for rainfall occurring in February to May / June, especially during the first half of the first rainfall season, could be explained. During the second half of the first rainfall season, when rivers flow, the negative effect is negligible. However, the positive correlation at the end of the year, occurring just after the peak of the second rainfall season, is in contrast with this reasoning. Possibly different breeding sites play a role at that time of the year. In the north and east, the climate is particularly dry from February – September. Here, rainfall during the middle of the year will provide the water required for mosquito breeding and humidity for survival, explaining the positive relationship found during the middle of the year, after the driest months. During the north-east monsoon (October – January) the rivers flow normally abundantly, and a negative correlation might be expected (based on the observations in the centre west of the country) during this period, but only for one district with poor quality malaria data (Mullaitivu) this was apparent, whereas for another close by district (Jaffna), the opposite was observed. There is no evidence that this mechanism plays a role in the north and east. Furthermore, in the north and east, the highest malaria peak is normally observed in January (just after the peak in the north-east monsoon rains). This is in line with the fact that early malariologists considered rivers not to be intimately involved in the mechanism of epidemic malaria in the dry zone [7]. The results from the analysis of correlation between Fisher transformed correlation coefficients and rainfall also suggested a different mechanism in the centre west from that in the rest of the country.

The fact that positive correlations were stronger at a lag of one month, and negative correlations were stronger at a longer lag of three months may be explained as follows: Within a one month period, rainfall can influence malaria transmission and cases positively by providing humidity which increases mosquito survival. One

month might not be long enough for rainfall to influence malaria cases through an effect on mosquito breeding. A negative effect of rainfall on mosquito breeding (for instance less than normal rainfall which might cause river pooling, which will have a delay in itself) will need a longer lag period to translate into a change in malaria cases.

Limitations of this study

This study was limited to linear rainfall – malaria relationships. For a better understanding of the biological mechanisms behind the observed relationships between rainfall and malaria cases, the link between rainfall and mosquito breeding and survival should be included. Long, high quality time series of entomological data were unfortunately not available for this purpose. Rainfall influenced variables, such as soil moisture saturation and river flow, are more directly linked to specific malaria vector breeding conditions. However, such variables are more expensive to measure and therefore often estimated using rainfall, and in the latter case using such a variable would have little advantages. There would be a clear advantage if for instance human interference with river flow, for purposes such as irrigation or power generation, could be taken into account. Such interference disrupts the relationship between rainfall and river flow, and hence the relationship between rainfall and malaria [119]. Apart from rainfall and rainfall related variables, another variable that is expected to have a strong temporal effect on malaria case count time series is malaria control intervention. This variable was not taken into account due to incomplete data. Also, control methods and insecticides have changed over time, making it a complex covariate. Temperature was not studied as it was considered of less importance than rainfall, showing little temporal variability (because Sri Lanka is situated close to the equator), and a large part of its temporal variability being governed by rainfall. Except for the hill country, situated in the centre of Sri Lanka, the temperature is conducive to malaria transmission throughout the year. Other environmental factors that are often considered in malaria studies are altitude and land use. These were not taken into account as these do not fluctuate (strongly) over time. Another limitation is the use of Gaussian models on transformed count data, whereas negative binomially distributed methods on untransformed data [120] may have been more appropriate. This study was performed on aggregated cases of *Plasmodium falciparum* and *P. vivax*. Although the seasonality of *P. vivax* is slightly less marked

than that of *P. falciparum*, possibly caused by relapses of *P. vivax* occurring well after infection, the seasonality is very similar [61]. In the current study, it was presumed that cases were infected in the district where they were recorded. Although with the large spatial units of districts only a small percentage of cases may have been acquired elsewhere, these would mostly be expected to have been acquired in neighbouring districts with similar rainfall patterns. Nevertheless, in districts with normally very low transmission such as Galle, Nuwara Eliya and Colombo, the proportion of cases from elsewhere might be much higher, and the relationships between rainfall and malaria for these districts should be interpreted with care.

Conclusions

Although malaria and rainfall showed high cross-correlations in many districts in Sri Lanka, variation from normal monthly malaria counts patterns showed limited cross-correlation with variation from normal monthly rainfall patterns, and therefore rainfall may have limited use for predicting malaria. Seasonally varying effects of rainfall on malaria case counts may explain weak cross-correlations in prewhitened series (as the cross-correlation analysis did not allow for a seasonally varying effect). There was a marked difference in the seasonally varying effect between the south-western quadrant and the rest of the country, which was probably related to differences in rainfall, but also to spatially different water requirements for optimum breeding conditions for the main malaria vector in Sri Lanka.

Author's contributions

OJTB conceptualized and conceived of the analysis, performed the data treatment and analysis, and drafted the manuscript. PV participated in the conceptualization, edited the manuscript and critically revised the statistical methodology. DMG participated in the conceptualization of the study and edited the manuscript. GNLG provided the data and helped define the scope of the paper. PHA participated in defining the approach to analysis, edited and critically reviewed the paper for intellectual content. All authors read and approved of the manuscript.

Acknowledgements

NOAA, NSF, EPA and EPRI Joint Program on Climate Variability and Human Health.

Chapter 7 – Models for short term malaria prediction in Sri Lanka

*This chapter was published in the Malaria Journal 2008, **7**:76*

Olivier JT Briët [*1,2], Penelope Vounatsou[2], Dissanayake M Gunawardena[3], Gawrie NL Galappaththy[4], Priyanie H Amerasinghe[5]

1 International Water Management Institute, P.O. Box 2075, Colombo, Sri Lanka, Telephone: +94112787404, Fax: +94112786854

2 Swiss Tropical Institute, Socinstrasse 57, P.O. Box CH-4002, Basel, Switzerland

3 US Agency for International Development, P.O. Box 7856, Kampala Uganda

4 Anti Malaria Campaign, Head Office Colombo, Sri Lanka

5 International Water Management Institute Sub Regional Office for South Asia, c/o ICRISAT, Patancheru, AP 502 324, Andhra Pradesh, India

* Corresponding author

Abstract

Background

Malaria in Sri Lanka is unstable and fluctuates in intensity both spatially and temporally. Although the case counts are dwindling at present, given the past history of resurgence of outbreaks despite effective control measures, the control programmes have to stay prepared. The availability of long time series of monitored/diagnosed malaria cases allows for the study of forecasting models, with an aim to developing a forecasting system which could assist in the efficient allocation of resources for malaria control.

Methods

Exponentially weighted moving average models, autoregressive integrated moving average (ARIMA) models with seasonal components, and seasonal multiplicative autoregressive integrated moving average (SARIMA) models were compared on monthly time series of district malaria cases for their ability to predict the number of malaria cases one to four months ahead. The addition of covariates such as the number of malaria cases in neighbouring districts or rainfall were assessed for their ability to improve prediction of selected (seasonal) ARIMA models.

Results

The best model for forecasting and the forecasting error depended strongly on the district. The addition of rainfall as a covariate improved prediction of selected (seasonal) ARIMA models modestly in some districts, but worsened prediction in other districts. Improvement by adding rainfall was more frequent at larger forecasting horizons.

Conclusions

Heterogeneity of patterns of malaria in Sri Lanka requires regionally specific prediction models. Prediction error was large at a minimum of 22% (for one of the districts) for one month ahead predictions. The modest improvement made in short term prediction by adding rainfall as a covariate to these prediction models may not be sufficient to merit investing in a forecasting system for which rainfall data are routinely processed.

Background

Malaria has been a major public health problem in Sri Lanka [2] until recently. Since the year 2000, incidence has dwindled [121] with only 591 reported cases for 2006 [122]. It is unstable and fluctuates in intensity both spatially and temporally, thus resources for control have to be spread in time and space to be prepared for outbreaks, which have occurred in the past despite very aggressive and effective malaria control operations [38]. Having a forecasting system in place will contribute to a more focussed approach for control, and have a positive impact on the resource allocation for malaria control over space and time. This paper explores different models for malaria case prediction, which is possible due to the availability of long, dense and reliable records of malaria cases and climate variables in Sri Lanka [26].

While many factors play a role in the spatial and temporal distribution of malaria, climate variability (both spatial variation of the long term seasonal mean of weather variables, and temporal aberrations from the long term seasonal mean) has been shown to be important in explaining its occurrence [9,109,123] and is considered a major determinant [124]. Temperature, rainfall, and humidity affect breeding and survival of a certain (sub) species of anopheline mosquitoes that carry the malaria parasite, as well as development of malaria parasites within vector mosquitoes, thereby creating a link between weather and malaria.

At present, there are no practical tools for temporal prediction of the occurrence of malaria based on observed rainfall or weather forecasts in Asia, although these are in development [125]. For Africa, such tools have been developed [94] and applied [126]. Recent work [127,128] focuses on malaria early warning systems, in which flags are raised when epidemics are expected. Setting the threshold for what is an epidemic (defined as a number of cases substantially exceeding that what is expected based on recent experience or what is thought normal) is subjective. The term 'epidemic' does not combine well with the term 'prediction' (if the expected number is predicted based on recent experience, the prediction can never be 'epidemic' according to the above definition). It is difficult to define, especially in Sri Lanka, at what level malaria incidence is thought to be normal, as the malaria time series show strong long-term fluctuations and it is, therefore, difficult to set thresholds. In general, disease forecasting is most useful to health services when it predicts case numbers two to six months ahead, allowing tactical responses to be made when

disease risk is predicted to increase (or decrease) [129]. For this reason this paper avoids the problem of setting epidemic thresholds, by focusing on forecasting malaria cases only.

Malaria case numbers are influenced by factors intrinsic to malaria such as infectivity, immunity and susceptibility of vectors and humans, and extrinsic, environmental factors such as rainfall. The number of possible models for malaria prediction is infinite. In biological process models, typically consisting of sets of equations, prediction can be done with details of all pathways, parameters and variables believed to be important for the dynamics of the disease [129]. In statistical models, temporal or spatial autoregressive terms account for the fact that case numbers depend on past or nearby case numbers through (sometimes cyclical) intrinsic processes, as well as for (unobserved) extrinsic auto correlated factors or factors with fading effects. This study was limited to some statistical models that are relatively easily implemented without taking into consideration complex biological processes and their parameters, and /or have been successful elsewhere in malaria forecasting studies. With sufficient temporal autocorrelation in malaria case time series, malaria cases can be predicted based on previous values [130]. However, predictions from statistical models are made under the assumption that the relationships established based on past observations remain the same into the future. Therefore, statistical models require experience with as wide a spectrum of conditions as possible. In this light, the present low case numbers, with a probable negative effect on immunity, have been unprecedented in the time series under study, and a caution should be in place. More complex statistical models can be constructed where malaria incidence in an area is, apart from its own previous values, also dependent on (previous) values in neighbouring areas, or covariates such as rainfall [107,131]. These latter models require more inputs and therefore more resource intensive to apply, particularly where covariate data need to be acquired and processed in a timely manner to be useful for forecasting. In this paper, it was examined which standard time series statistical model would be useful for forecasting malaria, and it was examined whether addition of rainfall to autoregressive models could improve malaria prediction in districts with one to four month forecasting horizons.

Methods

This section describes the data used for the analyses, methods for pre-processing of the data, types of models tested and the criteria for model selection.

Malaria data

The count of blood films examined for malaria as well as those positive for malaria per month reported by government health facilities and aggregated by medical officer of health (MOH) area (which represent sub district health administrative divisions) were provided by the Anti Malaria Campaign of Sri Lanka for the period 1972 – 2003. In addition, data aggregated by district were available for the years 2004 – 2005. For some of the records, the number of blood films examined was marked as "not received" (and therefore classified as missing). For 14.90% of the MOH area level records, the value was zero, or left blank. For the latter records, there was ambiguity as to whether the data value could be missing due to problems in data recording, or genuinely zero if no patients presented themselves for examination in that particular area in that particular month. As such, in a data cleaning procedure (see section on statistical methods), 1.4% of the records was declared as not available (NA) if the number of blood films examined was marked as "not received" (0.95%), or if the number of blood films could be classified as a lower additive outlier (0.44%). The data from districts in the north and east, where data gathering and reporting was affected by the armed conflict, had the largest percentage of data labelled not available: Jaffna (5.4%), Mannar (26.1%), Vavuniya (8.9%), Kilinochchi (2%), Trincomalee (2%) and Ampara (5.4%). After imputation, MOH area level data for positive cases were aggregated to district resolution and combined with the district level data (for the period 2004 – 2005).
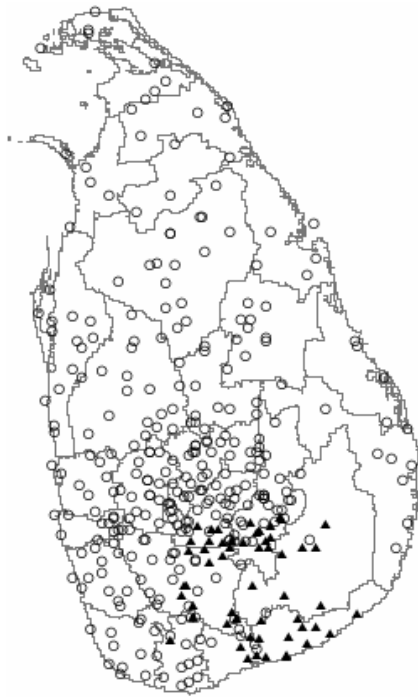
Figure 37 - Rainfall stations

Location of stations measuring rainfall for which monthly data (open circles) and daily data (solid triangles) were available. Grey lines represent current boundaries of the 25 districts. The time period for which data was available varied per station.

Rainfall data

Records of precipitation (rain fall) collected by 342 stations across the island were purchased from the Meteorological Department of Sri Lanka (see Figure 37). This consisted mostly of monthly aggregate data, but for an area in the south (Ruhuna), daily rainfall data were also available for 57 stations covering partly the districts of Ratnapura, Hambantota, Badulla and Moneragala, for the period January 1972 – March 2003. Three stations with consistently aberrant rainfall, detected through cross validation using kriging [99], were removed from the dataset. Monthly rainfall surfaces were created through spatial prediction using kriging [99]. From the daily data available, the monthly "rainy day index" was calculated for each station by dividing the number of days per month that rainfall was larger than zero by the number of days that a reading for rainfall was available. Monthly rainy day index surfaces were generated following the same procedure as for the total monthly rainfall. From each monthly rainfall surface, the average value of rainfall / rainy day index was extracted for each district.

Statistical methods

The monthly count of malaria positive blood slides in each district y_t were transformed to normality via the logarithmic transformation $z_t = \log(y_t + 1)$. The models tested included exponentially smoothing and auto-regressive integrated moving average (ARIMA) models [90]. As some of the district malaria count time series showed strong seasonality, seasonality was also modelled. In models using exponential smoothing, seasonality was included using the Holt-Winters procedure [90]. In ARIMA models, seasonality was included via three different approaches which are all widely used in literature: seasonality through fixed (monthly) effects; seasonality through harmonics; and through random effects using seasonal multiplicative auto-regressive integrated moving average (SARIMA) models. Whether or not covariates such as rainfall and concurrent malaria case counts in neighbouring areas improved the predictive ability of the models was also tested. In addition, the seasonal adjustment method used by Abeku and colleagues [130], was tested.

Exponentially weighted moving average models

The additive Holt-Winters prediction function (for time series with period length *s*) at time *t+h* is given by the following equation:

$$\hat{z}_{t+h} = m_{t,h} + S_{t,h}$$

where $m_{t,h}$ is the average number of cases at time *t+h* expressed as a trend $r_{t,h}$ and an overall mean term a_t, that is $m_{t,h} = r_{t,h} + a_t$. $S_{t,h}$ is a seasonality term at time *t+h*, such that $S_{t,h} = S_{t-s+1+(h-1)\bmod s}$ where $(h-1)\bmod s$ is the remainder of *h-1* after division by *s* (e.g. $14 \bmod 12 = 2$). Thus

$$\hat{z}_{t+h} = a_t + hr_t + S_{t-s+1+(h-1)\bmod s} \qquad\qquad \textbf{(1)}$$

where a_t, r_t and S_t are calculated by the following recursive functions:

$$a_t = \alpha(z_t - S_{t-s}) + (1-\alpha)(a_{t-1} + r_{t-1});$$

$$r_t = \beta(a_t - a_{t-1}) + (1-\beta)r_{t-1};$$

$$S_t = \gamma(z_t - a_t) + (1-\gamma)S_{t-s}.$$

Both seasonal and non-seasonal (with γ fixed to 0) models were tested using the function "HoltWinters" in the package "stats" of the statistical software package "R".

(S)ARIMA regression models

It was assumed that z_t is Gaussian distributed, $z_t \sim N(\mu_t, \sigma^2)$, with mean μ_t and variance σ^2. Further, it was assumed that

$$\mu_t = f(z_t, d, p, x_t) + g(u_t, q) \qquad (2)$$

where $f(z_t, d, p, x_t)$ and $g(u_t, q)$ model the temporal correlation as

$$f(z_t, d, p, x_t) = \Phi_p(B)(1-B)^d (x_t - z_t) + z_t \text{ and } g(u_t, q) = \Theta_q(B)u_t - u_t$$

where

$$\Phi_p(B) = 1 - \phi_1 B - \ldots - \phi_p B^p;$$

$$\Theta_q(B) = 1 - \theta_1 B - \ldots - \theta_q B^q;$$

u_t is Gaussian white noise;

$x_t = m_t + S_t;$

S_t models the seasonal process;

m_t models the mean of z_t

B is a backshift operator with $B^d(z_t) = z_{t-d}$.

The seasonality in the ARIMA models of equation 2 was modelled by fixed effects. In particular it was assumed:

- $S_t = 0$ (A non seasonal model),

- $S_t = \sum_{k=1}^{k=12} (\alpha_k \delta_{k,t})$ where $\delta_{k,t} = \begin{cases} 1 \text{ if } t = \text{nk} \\ 0 \text{ if } t \neq \text{nk} \end{cases}$, $n = \{1, 2, \ldots\}$

(Seasonality through fixed effects for months: Note that in this model m_t does not contain an intercept to avoid over parameterisation),

- $S_t = \sum_{i}^{2} A_i \sin(2\pi f_i t + 2\pi\varphi_i)$

a second order harmonic component where A_i is the amplitude of harmonic i; f_i is the frequency of harmonic i, with $f_1 = 1/s$, $f_2 = 2/s$; and φ_i is the phase shift (in units of time) of harmonic i.

Also, a multiplicative seasonal ARIMA$(p,d,q)*(P,D,Q)$ model (henceforth SARIMA) was considered with period s, obtained by modifying equation 2 into

$$\mu_t = f(z_t, d, p, D, P, s, m_t) + g(u_t, q, Q, s)$$

where

$$f(z_t, d, p, D, P, s, m_t) = \Phi_p(B)(1-B)^d (1-B^s)^D \Phi_P^*(B^s)(m_t - z_t) + z_t;$$

$$g(u_t, q, Q, s) = \Theta_q(B)\Theta_Q^*(B^s)u_t - u_t;$$

$$\Phi_P^*(B^s) = 1 - \phi_1^* B^s - \ldots - \phi_P^* B^{sP};$$

$$\Theta_Q^*(B^s) = 1 - \theta_1^* B^s - \ldots - \theta_Q^* B^{sQ};$$

and $\Phi_p(B)$, $\Theta_q(B)$, u_t, m_t and B as explained above.

The function "arima" in the package "stats" of the statistical software "R" was used to calculate the prediction criterion. Tested models included all (Gaussian) ARIMA models possible with combinations of parameters (p, d, q) with $p, q \in \{0,1,2\}$ and with $d = 1$, without explanatory variables, and all (Gaussian) SARIMA models possible with combinations of parameters (p, d, q, P, D, Q) with $p, q \in \{0,1,2\}$ and $d = 1$ and $P, D, Q \in \{0,1\}$, also without explanatory variables. An intercept was not included in the mean as it drops out of the equation due to differencing ($d = 1$). The differencing also removes effects of trends such as potentially caused by population growth.

Covariates were included in the term m_t. In particular, 1) $m_t = \sum_j \beta_j z_{j,t-1}$ where $z_{j,t-1}$ is the transformed malaria count at month t - 1 in neighbour j; 2) $m_t = \beta \chi_{t-l}$ where χ_t is the rainfall parameter in month t - l with l = lag. Rainfall was considered at lags of one to four months preceding malaria and in the following forms: untransformed monthly rainfall, logarithmically transformed monthly rainfall, rainy day index (for those districts with daily rainfall available), monthly rainfall factored into quintiles (in case of non-linear relationships), and rainfall with a separate coefficient for each of the twelve calendar months, *i.e.* a coefficient for January rainfall, one for February, etc., in order to allow for seasonally varying effects. For each district, covariates were tested by including them into the (S)ARIMA model that performed best for the respective district and lag.

Estimation of non-available malaria count data

In a data cleaning procedure, the time series of blood film counts in MOH areas were logarithmically transformed to normality (after the value one was added to the data). Under the null hypothesis, each observation was assumed to be part of a seasonal autoregressive integrated moving average (SARIMA) process with parameters $p = 0$, $d = 1$, $q = 1$, $P = 0$, $D = 1$, and $Q = 1$. Observations were marked as additive outlier if the likelihood ratio test statistic (for an additive outlier) for the observation was below a threshold of -6 [98], and were classified as not available. For those observations classified as not available that were not at the beginning or end of a series, values for the number of malaria positive blood films were estimated through a one-step-ahead SARIMA forecasting model on the original series and on the reversed series, and the two estimates were averaged. This approach has been discussed by Mwaniki and colleagues [101]. Finally, the MOH area data series were aggregated to district resolution before analysis, as these spatial units remained constant over the study period, whereas for many MOH areas boundaries changed (within district boundaries) over the study period.

Seasonal adjustment method with last three observations

Abeku and colleagues [130] tested a seasonal adjustment method on malaria data in Ethiopia and found that it performed better in comparison to SARIMA models. They obtained best results when using a three year "training" time series. The prediction formula used is as follows:

$$\hat{z}_t = \tfrac{1}{3}\sum_{k=1}^{k=3}\left\{z_{t-12k}\right\} + \tfrac{1}{3}\sum_{l=1}^{l=3}\left\{z_{t-l} - \tfrac{1}{3}\sum_{k=1}^{k=3}\left\{z_{t-12k-l}\right\}\right\}.$$

Model evaluation

For each district, model parameters were estimated on approximately the first half of the malaria case time series (January 1972 – December 1987), and one to four step ahead (out of sample) predictions were made on the second half (January 1988 – December 2005) with the parameters fixed.

For selection of the best predictive models, all models tested were evaluated on the prediction criterion which was defined as the mean absolute relative error (*mare*) of back transformed out of sample predictions:

$$mare = \sum_{i=1}^{N} \left| \frac{y_t - \hat{y}_t}{y_t + 1} \right| \Big/ N$$

where $\hat{y}_t$ is the predicted number of malaria positive cases at time t, and N is the number of predictions. Predictions needed to be genuinely out of sample in order to prevent bias towards more highly parameterized models. The *mare* was used rather than mean square error, as the malaria count time series show widely differing variances across the series [90]. The best model was that with the lowest prediction criterion for a given time series.
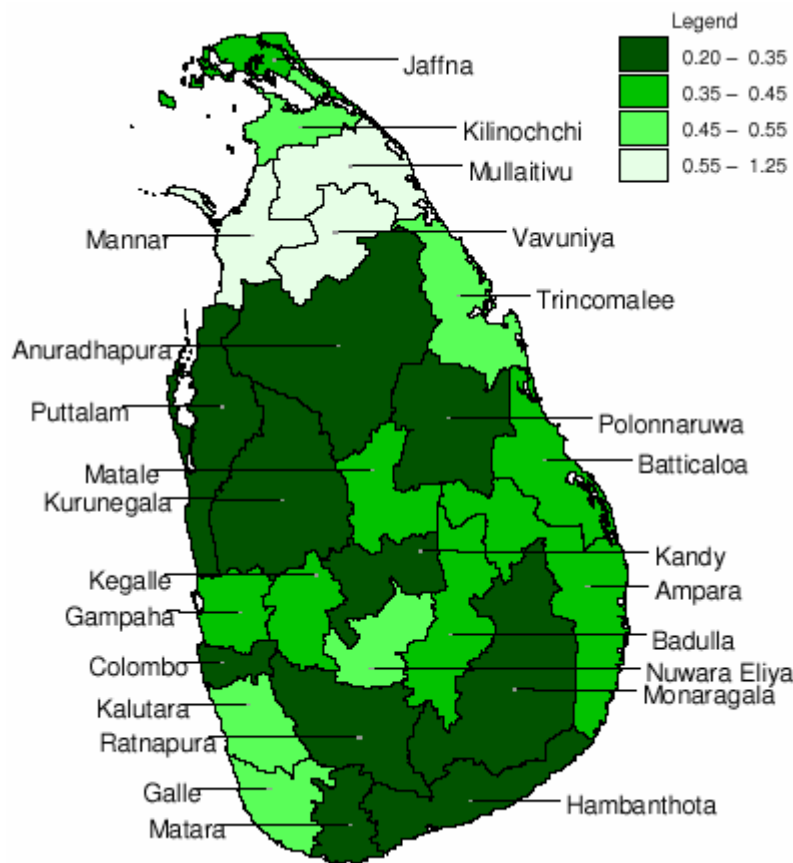
Figure 38 - Mean absolute relative error in districts at a 1 month forecasting horizon

Mean relative absolute error of out of series prediction at a forecasting horizon of 1 month ahead for districts in Sri Lanka for the best model (without the inclusion of rainfall as a covariate) tested.

Table 8 - Mean absolute relative error of out of series prediction at forecasting horizons of 1 to 4 months ahead for districts in Sri Lanka for the best (S)ARIMA model tested.

District	Horizon 1		Horizon 2		Horizon 3		Horizon 4	
	Criterion	Model (*pdqPDQ*)	Criterion	Model (*pdqPDQ*)	Criterion	Model (*pdqPDQ*)	Criterion	Model (*pdqPDQ*)
Ampara	0.37	012SOH	0.48	012101	0.58	012101	0.60	012101
Anuradhapura	0.23	211101	0.37	210110	0.45	012110	0.51	210110
Badulla	0.43	110SOH	0.62	111SOH	0.75	212101	0.74	112100
Batticaloa	0.36	010011	0.54	012101	0.66	012101	0.78	012101
Colombo	0.35	011000	0.38	112000	0.43	211001	0.46	011000
Galle	0.49	212002	0.58	211101	0.63	211101	0.71	211110
Gampaha	0.40	011111	0.56	011SOH	0.67	011SOH	0.78	011SOH
Hambantota	0.31	010101	0.47	110101	0.60	210101	0.71	210101
Jaffna	0.42	010011	0.58	012111	0.74	012011	0.82	012SOH
Kalutara	0.54	112100	0.72	011000	0.79	110000	0.79	110000
Kandy	0.33	012101	0.43	012101	0.48	112SOH	0.51	212SOH
Kegalle	0.37	010SOH	0.55	211011	0.66	211SOH	0.75	211SOH
Kilinochchi	0.51	010101	0.95	010101	2.13	111010	2.13	010002
Kurunegala	0.25	011011	0.41	010011	0.53	011011	0.63	011011
Mannar	1.16	011100	0.97	012101	1.10	112100	1.18	111101
Matale	0.37	110101	0.53	110101	0.62	212011	0.70	112011
Matara	0.35	212101	0.40	011101	0.46	212101	0.49	0110111
Moneragala	0.29	110100	0.40	011100	0.48	210100	0.56	011100
Mullaitivu	1.03	111100	1.70	112000	2.00	110000	2.58	111SOH
Nuwara Eliya	0.48	212111	0.58	212101	0.66	212101	0.68	111000
Polonnaruwa	0.32	111101	0.47	012101	0.57	111011	0.66	111011
Puttalam	0.35	010101	0.46	010101	0.60	212101	0.72	010101
Ratnapura	0.30	011111	0.43	012111	0.50	210111	0.57	112111
Trincomalee	0.53	112000	0.79	010100	1.05	010100	1.15	112111
Vavuniya	1.22	012000	1.43	012101	1.41	211101	1.48	012101

Legend: *pdq* = order of autoregressive component, integrated component and moving average component; *PDQ* = order of seasonal autoregressive component, seasonal integrated component and seasonal moving average component; SOH = seasonality through second order harmonic;

Results

The best model (without extrinsic explanatory variables) varied by district and forecasting horizon (Table 8). For instance, for the district of Ampara, for a one month forecasting horizon, the best model was an ARIMA (2,1,1) model with seasonality modelled through a harmonic with a period of one year and a harmonic with a period of six months. For further forecasting horizons, the ARIMA(0,1,2) model with seasonality through a first order seasonal autoregressive and a first order seasonal moving average component was best for the district of Ampara. The best model was most often of the SARIMA class, followed by the class of models modelling seasonality through second order harmonics. For a few districts, at some forecasting horizons, exponential smoothing was best (Table 9). The seasonal adjustment method performed worst (Not shown). The mean relative absolute error of forecasts varied over the districts (for the same forecasting horizon, see Figure 38), and increased with forecasting horizon. The *mare* was relatively high for the districts Galle and Kalutara in the south west, and Nuwara Eliya in the central hill country, which have low malaria endemicity. The *mare* was also (very) high for the districts affected by the armed conflict in the north and east. Within a class of models, the most complicated model tested was not necessarily the best model, and often the prediction improvement obtained by fitting an extra (S)ARIMA parameter as compared to more parsimonious models was marginal.

In the analysis of covariates for the mean term, only the (S)ARIMA models shown in Table 8 were tested. Only for the districts Mannar and Ampara, inclusion of malaria in neighbouring districts lowered one month ahead *mare*, with 6.8% and 4.6%, respectively. For many other districts, inclusion of malaria in neighbouring districts raised the *mare*. Inclusion of a rainfall variable lowered the *mare* with 2.5% or more for eight districts at one or more horizons (Table 10). Logarithmically transformed rainfall lowered the *mare* for Gampaha District (at horizons of three and four months), Mannar District (at a horizon of one month) and Vavuniya District (at a horizon of four months). Logarithmically transformed rainfall with a separate coefficient for each calendar month lowered the *mare* for Ratnapura District (at horizons of three and four months), and Trincomalee District (at horizons of two and three months). Rainfall factored into quintiles (allowing a non-linear relationship) lowered the *mare* for Moneragala District (at a horizon of two months) and Mullaitivu District (at a

horizon of one month). The rainy day index lowered the *mare* for Moneragala District (at a horizon of three months), and the rainy day index with a separate coefficient for each calendar month lowered the *mare* for Badulla District at a horizon of four months.

Table 9 - Mean absolute relative error of out of series prediction at forecasting horizons of 1 to 4 months ahead for districts in Sri Lanka for Holt Winters models.

| District | Horizon 1 | | Horizon 2 | | Horizon 3 | | Horizon 4 | |
Model	H	HW	H	HW	H	HW	H	HW
Ampara	0.43	0.39	0.65	0.52	0.83	0.63	0.86	0.67
Anuradhapura	0.34	*0.22*	0.66	*0.35*	0.99	*0.45*	1.22	0.53
Badulla	0.46	0.54	0.67	0.75	0.87	0.95	0.84	0.96
Batticaloa	0.41	0.41	0.65	0.65	0.82	0.82	0.97	0.97
Colombo	*0.35*	0.37	0.39	0.43	0.44	0.48	*0.46*	0.53
Galle	0.50	0.61	0.59	0.74	0.67	0.83	0.79	0.96
Gampaha	0.43	0.43	0.59	0.59	0.70	0.70	0.78	0.78
Hambantota	0.36	0.36	0.57	0.56	0.76	0.73	0.88	0.87
Jaffna	0.43	0.46	0.62	0.63	0.79	0.85	0.85	0.97
Kalutara	0.55	0.61	0.72	0.80	0.81	0.91	0.88	0.97
Kandy	0.37	0.37	0.50	0.50	0.56	0.57	0.57	0.57
Kegalle	0.39	0.40	0.63	0.62	0.83	0.82	0.94	0.95
Kilinochchi	0.58	0.60	1.08	1.12	2.50	2.26	2.70	2.17
Kurunegala	0.34	0.26	0.61	0.43	0.76	0.57	0.85	0.70
Mannar	1.41	1.57	1.74	1.98	1.61	2.63	1.78	2.28
Matale	0.45	0.41	0.73	0.63	0.96	0.74	1.13	0.81
Matara	0.37	0.35	0.42	0.40	0.49	0.48	0.52	0.52
Moneragala	0.31	0.31	0.42	0.41	0.54	0.52	0.62	0.63
Mullaitivu	1.08	1.19	1.73	1.70	2.21	2.54	2.73	*2.38*
Nuwara Eliya	0.49	0.50	0.61	0.60	0.69	0.69	0.69	0.69
Polonnaruwa	0.37	0.37	0.60	0.60	0.76	0.76	0.82	0.82
Puttalam	0.42	0.37	0.67	0.49	0.88	0.64	1.00	0.76
Ratnapura	0.36	0.31	0.55	0.47	0.64	0.56	0.74	0.66
Trincomalee	0.53	0.56	0.82	*0.75*	1.15	*0.97*	1.35	*1.07*
Vavuniya	1.89	2.02	2.82	3.93	2.45	14.21	2.19	4.11

H = Holt's two parameter exponential smoothing; HW = Holt-Winters' three parameter exponential smoothing (including seasonality). Values in bold italic represent a better mare as compared to the best (S)ARIMA model (without rainfall).

Table 10 - Districts in Sri Lanka for which inclusion of a covariate in the mean term of the best (S)ARIMA model tested improved the mean absolute relative error of out of series prediction at forecasting horizons of 1 to 4 months ahead.

District	Horizon (months)	Lag (months)	covariate	Improve-ment (%)
Badulla	4	4	rainy day index, with a separate coefficient for each calendar month	6.5
Gampaha	3	4	logarithmically transformed total monthly rainfall (mm)	3.8
Gampaha	4	4	logarithmically transformed total monthly rainfall (mm)	4.5
Mannar	1	2	logarithmically transformed total monthly rainfall (mm)	5.2
Moneragala	2	2	monthly rainfall factored into quintiles	4.1
Moneragala	2	3	rainy day index	4.6
Moneragala	3	3	rainy day index	3.2
Mullaitivu	1	1	monthly rainfall factored into quintiles	2.6
Ratnapura	3	4	logarithmically transformed total monthly rainfall (mm), with a separate coefficient for each calendar month	3.9
Ratnapura	4	4	logarithmically transformed total monthly rainfall (mm), with a separate coefficient for each calendar month	3.6
Trincomalee	2	2	logarithmically transformed total monthly rainfall (mm), with a separate coefficient for each calendar month	8.4
Trincomalee	3	3	logarithmically transformed total monthly rainfall (mm), with a separate coefficient for each calendar month	9.2
Vavuniya	4	4	logarithmically transformed total monthly rainfall (mm)	2.5

Discussion

Models without extrinsic explanatory variables

The mean relative prediction error calculated by the best models (without extrinsic explanatory variables) tested was quite large for many districts. However, the models were fitted to only half of the length of the time-series available, because for the purpose of model testing out of sample predictions are required. It is expected that for application in a forecasting system where the full series are used for fitting, the error will be reduced. For some districts in the north, the forecasting error was particularly large. For these districts, a relatively large proportion of the data had been imputed, and the quality of the existing data is likely compromised by the armed conflict in this region. General issues related to data quality are discussed elsewhere [61]. In this (primarily) temporal study, issues relating to spatial variation in data quality (*e.g.* through access to health facilities for sampling) are of less importance than those

pertaining to temporal aspects, such as the deployment of mobile clinics during specific periods. Despite the sometimes large prediction errors, especially for larger forecasting horizons, prediction intervals yielded by these models (albeit less accurate for low predicted mean counts due to the Gaussian approximation used) could aid the AMC in assessing the risk of malaria in the near future, and adjust resources for preparedness accordingly. Although the best model selected varied among districts and over forecasting horizons, the difference between models was often small. Instead of specifying a different model for each situation, for practical implementation, it may be worth selecting for each district (or even group of districts) one model that performs well on average over a range of forecasting horizons (and districts within the group), provided that the prediction quality does not deteriorate more than a set percentage. A pilot forecasting system using district specific SARIMA models is currently being tested by the AMC in Sri Lanka (the system uses currently models without explanatory variables because a system to incorporate newly observed explanatory data is not yet in place). As the spatial resolution of the forecasting models presented is at district level, predictions will not help spatially targeted control at sub district level. For this, regional malaria control officers will have to rely on their experience of where within the district cases tend to occur if they occur, possibly aided by existing malaria maps at sub district level [61].

Models including rainfall as explanatory variable

It should be kept in mind that the malaria count data are not a direct proxy of new malaria infections or even infective bites. Recrudescent and relapsing cases (mostly due to ineffective drugs) occur multiple times in the malaria dataset, and immunity may play a role during periods of higher endemicity, thus weakening explanatory power of a variable such as rainfall, which would probably be better at explaining infectious bites [132].

Rainfall improved the prediction for eight districts, at one or more of the tested forecasting horizons, but it also worsened the prediction for other districts at various horizons. For some districts, the rainfall preceding malaria by two months improved performance at forecasting horizons of three and four months (not shown), but this is of little use in a forecasting system, unless rainfall can be predicted with high confidence (one to two months ahead). Although only tested for four districts, the rainy day index was promising as it improved results for two out of these.

The results of the analysis incorporating rainfall was remarkably different from what would be expected based on results of cross correlation analysis of malaria and rainfall. Although both (pre-whitened) cross correlations [96] and the improvements of rainfall to prediction accuracy were generally low, (pre-whitened) cross correlation analysis indicated that prediction accuracy for a number of districts in the centre-west and centre-south of the country was likely to benefit from rainfall (with a single, negative coefficient), and the present analysis found such improvement only for Gampaha District.

Varying seasonal effects of rainfall on malaria in seasonal inter-annual analysis [96] suggested that especially for districts in the north and east, prediction models would benefit from addition of rainfall with a monthly varying coefficient. However, only for the districts Trincomalee and Ratnapura (the latter situated in the south), this was the case. It is possible that for most of the districts, the training time series was too short (sixteen years) to estimate each of the twelve coefficients reliably.

For most districts (except the districts Mannar, Mullaitivu and Vavuniya), presumably due to strong short term temporal auto correlation, the observed number of malaria cases is a good predictor at a one month forecasting horizon (Figure 38). Interestingly, for Mannar and Mullaitivu, rainfall could improve predictions somewhat at this horizon. For the remaining six districts for which rainfall was found to improve predictions, this was mostly at larger forecasting horizons. For prediction with an even further forecasting horizon, use of predicted rainfall might be feasible. Rainfall could be predicted up to several months ahead using, for instance, the El Niño southern oscillation index, which in itself has proven to have a statistically significant relationship with malaria epidemics in the pre control era [9]. During periods when case reporting might be compromised (e.g. through civil war), rainfall may gain in relative importance as a predictor of the true number of cases, although this would be difficult to validate. It should be noted that high cross correlations between (not prewhitened, seasonal) malaria time series and extrinsic covariates (with seasonality) such as rainfall and temperature may exist [96], and these covariates may perform well in models that do not include fixed seasonal or auto-regressive seasonal terms, but not necessarily better than models that do include fixed seasonal or auto-regressive seasonal terms.

The contribution of rainfall to reducing the prediction error was modest, and it is arguable whether or not the achieved reduction in the prediction error (for instance, a reduction of 9.2%, bringing the mean relative error down from 1.05 to 0.95 at a horizon of three months in the district Trincomalee) merits investing in a system where at the end of each month, monitored rainfall data are collected and processed to obtain average values for the district surface for entering into the prediction model. However, as such data become increasingly available with the development of monthly rainfall data monitoring systems at district and sub district scale [133], the reduction in prediction error gained might well outweigh the (reduced) effort.

The modest improvement found by including rainfall compares well to a study in Ethiopia [131]. That study also found modest improvement (a reduction of 10.7% in the number of one step ahead predictions that were more than twice or less than half the observed value) by a model incorporating rainfall and temperature over a simple model using the previous value as predictor (an ARIMA(0,1,0) model). However, the improvement would probably have been less if it had been tested on out of sample predictions. It is interesting to note that in this same paper, Abeku and colleagues [131] found that the ARIMA(0,1,0) performed better than the seasonal adjustment method (which performed unsatisfactorily on Sri Lankan data), whereas previously the seasonal adjustment method had been reported as superior [130]. Another study carried out in Ethiopia found that a prediction model including rainfall, temperature and cases in the previous month performed well in flagging potential epidemics as compared to observed cases, but was not compared with a prediction model based on past cases alone [127], therefore, the relative improvement cannot be assessed and compared to the improvement found in the present study.

Some explanatory variables not considered in this study and suggestions for future work

Other rainfall related variables such as soil moisture content and river flow might give better results than observed rainfall, as these are more directly related to malaria vector breeding conditions. The main malaria vector in Sri Lanka, *Anopheles culicificacies*, primarily breeds in pooling rivers (although there are also other habitats such as those mentioned above). In general, rainfall and river flow show strong cross correlations, and therefore rainfall can serve as a proxy for river flow. However, under dryer conditions –important for vector breeding– direct runoff will be decreased

and river flow will be increasingly influenced by other hydrological processes such as percolation and evapo-transpiration. As river flow and soil moisture are difficult to measure over large surfaces (and long term data with good spatial resolution are not readily available), they can be estimated through modelling, although rainfall will remain an important variable in such models [134]. Vegetation indices such as the normalized difference vegetation index (NDVI) were not studied because it can be argued that in a primarily temporal analysis, observed station rainfall is more closely related to breeding conditions than temporal changes in NDVI. Furthermore, remote sensing images of Sri Lanka suffer from obstruction by cloud cover during the monsoons. Apart from rainfall and rainfall related variables, another variable that is expected to have a strong temporal effect on malaria is vector control. This variable was not taken into account due to incomplete data. The effect of rainfall on malaria could be studied for data in the pre-control era [9], but with the reality of control measures being carried out this will be of little use for current prediction. The primary goal of this study was to develop an easy to use practical prediction model with the data available. The effect of missing important variables in a prediction model will not invalidate results, but will be expressed in larger prediction intervals. Nevertheless, inclusion of a malaria control variable, at least for those districts with relatively complete data, merits investigation.

Temperature was not considered as part of the present analysis, based on the assumption that it was of less importance than rainfall, showing little temporal variability (because Sri Lanka is situated close to the equator), and a large part of its temporal variability being governed by rainfall. Except for the hill country (covering largely the districts Nuwara Eliya, Kandy and Badulla), the temperature is generally conducive to malaria transmission throughout the year. However, for parts of the districts Kandy and Badulla, temperature could be limiting during part of (but not the whole of) the year and its role in these districts merits further investigation. In a study on the relationship between malaria and rainfall and temperature in Ethiopia, rainfall was found to be important in hot districts which were situated below 1650 m, but not in cold districts above 1650 m, where minimum temperature was significant [105]. Other environmental factors that are often considered in malaria studies are altitude and land use. Altitude was not taken into account because it does not fluctuate over time, and is thus of no use for temporal forecasting. Land use databases with monthly

temporal resolution were not available. The performance of rainfall over several lag times accumulated, or at thresholds, remains to be explored. It is tempting to try to build a space-time malaria forecasting model for the whole of Sri Lanka with more statistical power than a separate model for each district. Such a model should allow for regionally varying functions of covariates and take into account spatial auto correlation between districts or even MOH areas. For the purpose of a forecasting system, with the presently small malaria counts in most districts, it might be more appropriate to model the malaria counts directly following a negative binomial distribution, rather than through a logarithmic transformation [120]. Although easy to implement (and used in this work as well as work by others), the Gaussian (Normal) approximation of the malaria count data (after logarithmic transformation) is likely to affect the accuracy of the predictions, particularly when the counts are expected to be close to zero. Most importantly, this may yield inaccurate prediction intervals during periods of low case numbers. Although not so important in the current study (where models were evaluated on the means of the predictions), prediction intervals are very important for control planning.

Author's contributions

OJTB conceptualized and conceived of the analysis, performed the data treatment and analysis, and drafted the manuscript. PV participated in the conceptualization, edited the manuscript and critically revised the statistical methodology. DMG participated in the conceptualization of the study and edited the manuscript. GNLG provided the data and helped define the scope of the paper. PHA participated in defining the approach to analysis, edited and critically reviewed the paper for intellectual content. All authors read and approved of the manuscript.

Addendum to Chapter 7 – Implementation of a malaria forecasting system

Introduction

One of the main objectives of the project "Analysis of impacts of climate variability on malaria transmission in Sri Lanka and the development of an early warning system" was to provide the Anti Malaria Campaign of Sri Lanka with a malaria forecasting system. In this addendum, it was described how the models and findings described in Chapter 7 were used to implement a forecasting system.

The introduction of a malaria forecasting system was done in two phases. The first system was introduced in June 2006, and feedback from RMOs was obtained during a workshop held at IWMI on the 30th of August 2006. Based on the feedback and further analysis, a multi-model system was developed to suit different districts. The new models were shared during a workshop held at IWMI in September 2007, where the RMOs and AMC staff were introduced to software. This software allowed RMOs to make the predictions in their district. Subsequently, the AMC headquarters agreed to make all predictions every month, as not all the RMOs did have computer facilities.

Forecasting model selection

The models in the software provided for the malaria forecasting system were all of the SARIMA class, and did not incorporate external climate variables such as rainfall.

Although observed rainfall could improve prediction in certain districts (Chapter 7), it was not included in the model. The reason for this is that (at the time of writing), there was no system operational which provides the Anti Malaria Campaign with up to date monthly district rainfall totals[c] at the end of each calendar month, or even predicted into the future. A model which depends on such data would thus not be functional in a forecasting system.

---

[c] Monthly district rainfall totals were used during the testing of the capacity of rainfall to improve predictions, as these were deemed to be the best measure of rainfall over the entire district. The monthly district rainfall totals were calculated by interpolating gauge data from 342 stations across the island using kriging, and extracting the total value for the surface of a district. As data acquisition of such a large number of stations might be costly and cause delays, a system whereby some important main rainfall stations are used, together with established statistical correlations with other stations, might be feasible, but this would require some programming and testing.

Table 11 - Quality of out of series predictions at forecasting horizons of 1 to 2 months ahead made with the SARIMA models recommended for districts in Sri Lanka for the most recent 24 month period (July 2006 – July 2007).

District	SARIMA (pdqPDQ)	Horizon 1				Horizon 2			
		% mare	mw 95%pi	% below	% higher	% mare	mw 95%pi	% below	% higher
Ampara	012101	55.0	5.4	8.3	0.0	75.0	7.9	4.2	4.2
Anuradhapura	210110	60.8	20.9	4.2	12.5	90.9	41.0	4.2	8.3
Badulla	112100	6.3	1.9	0.0	0.0	12.5	2.5	0.0	0.0
Batticaloa	012101	70.5	4.7	12.5	8.3	90.8	8.8	8.3	4.2
Colombo	211001	20.8	1.9	0.0	0.0	27.1	3.0	4.2	0.0
Galle	211101	0.0	2.8	0.0	0.0	0.0	3.0	0.0	0.0
Gampaha	011111	105.1	8.0	4.2	8.3	137.9	11.6	8.3	0.0
Hambantota	210101	61.0	3.5	4.2	12.5	101.8	5.0	8.3	8.3
Jaffna	012111	58.0	3.1	4.2	4.2	43.4	5.2	0.0	0.0
Kalutara	110000	31.9	3.7	0.0	0.0	20.1	4.6	0.0	0.0
Kandy	012101	35.4	1.9	0.0	4.2	45.8	2.6	0.0	4.2
Kegalle	211011	22.2	1.9	0.0	0.0	27.8	3.2	0.0	0.0
Kilinochchi	010101	6.3	3.2	0.0	0.0	22.9	6.6	0.0	0.0
Kurunegala	011011	74.0	26.1	16.7	12.5	79.7	41.5	8.3	4.2
Mannar	012101	10.4	5.5	0.0	0.0	8.3	8.1	0.0	0.0
Matale	112011	36.7	2.3	4.2	8.3	38.1	3.1	0.0	8.3
Matara	011101	31.9	2.1	0.0	4.2	34.7	2.3	0.0	0.0
Moneragala	210100	24.3	1.8	0.0	0.0	21.5	2.4	0.0	4.2
Mullaitivu	112000	8.7	5.7	0.0	0.0	13.0	8.7	0.0	0.0
Nuwara Eliya	212101	0.0	2.0	0.0	0.0	0.0	2.1	0.0	0.0
Polonnaruwa	111011	35.0	2.9	0.0	4.2	45.4	4.5	4.2	4.2
Puttalam	010101	37.5	5.9	0.0	4.2	64.7	10.9	0.0	0.0
Ratnapura	012111	25.7	2.0	0.0	4.2	31.3	2.8	0.0	4.2
Trincomalee	112000	80.2	54.6	0.0	4.2	157.6	114.5	0.0	0.0
Vavuniya	012101	80.9	29.3	0.0	8.3	126.3	36.9	0.0	8.3

Legend: pdq = order of autoregressive component, integrated component and moving average component; PDQ = order of seasonal autoregressive component, seasonal integrated component and seasonal moving average component; %mare = mean relative absolute error of predictions expressed as percentage; mw95%pi = mean width of 95% prediction interval; %below = percentage of observations lower than lower 95% prediction interval boundary; %higher = percentage of observations higher than upper 95% prediction interval boundary; horizon = the number of month predictions are made ahead.

The models provided assume that the outcome variable (logarithmically transformed malaria case count data) is Gaussian distributed. This assumption might be violated, especially when case counts are close to zero. This could yield incorrect prediction intervals.

Model performance

For the most recent two years of data available, for each district, predictions at a one and a two month forecasting horizon were made using the recommended models and compared to observed cases (Table 11). The mean absolute relative prediction error (*mare*) is a good measure of model accuracy. The results for the *mare* were very variable, depending on the district. This was due, in part, to the extremely low number of malaria cases over the most recent two year period of the data. For instance, in the district of Galle, no cases were observed, and predictions were zero for all months, hence the extremely low *mare* found of 0 %. The mean width of the 95% prediction interval (*mw95%pi*) is a measure of model precision. The *mw95%pi* also varied largely over the districts. The *mw95%pi* was particularly large for districts for which one or more of the observations were outside the 95 % prediction interval. Although 5% of the observations are expected to lie outside the 95 % prediction interval, it was striking that for some districts, this percentage was higher (up to 29% in Kurunegala district). The example of the district of Kurunegala (Table 12) shows that after one observation occurred outside the 95 % prediction interval, prediction intervals of shortly following predictions were much wider. The 95% prediction interval in the models used might be too narrow for some districts and not account for "erratic" behaviour of malaria positive counts time series. This is possibly due to the fact that predictions were made using Gaussian models on logarithmically transformed data, which may be less accurate at low malaria count levels.

Forecasting system - Phase 1: Predictions at one and two months' forecasting horizons

The AMC head office provided IWMI with the most recent monthly district malaria positive case counts through E-mail or telephone communications. The AMC head office obtained these data from RMOs at monthly held review meetings, often held towards the end of the month. At IWMI, the new monthly district malaria case data were entered manually and appended to the data file. Consequently, predictions at a

Table 12 - Observed values and predictions at a forecasting horizon of 1 month ahead for Kurunegala District

Month	Observed	Prediction	L95%pib	U95%pib	w95%pi	%re	below	higher
Aug-05	27	12	5	24	19	-54	0	1
Sep-05	21	28	13	58	45	32	0	0
Oct-05	21	21	10	43	33	0	0	0
Nov-05	24	23	11	48	37	-4	0	0
Dec-05	47	33	16	68	52	-29	0	0
Jan-06	96	54	26	109	83	-43	0	0
Feb-06	34	69	34	142	108	100	1	0
Mar-06	15	20	9	41	32	31	0	0
Apr-06	11	9	4	19	15	-17	0	0
May-06	10	14	6	29	23	36	0	0
Jun-06	8	12	6	26	20	44	0	0
Jul-06	25	9	4	18	14	-62	0	1
Aug-06	10	27	13	57	44	155	1	0
Sep-06	16	8	4	18	14	-47	0	0
Oct-06	13	17	8	35	27	29	0	0
Nov-06	7	14	6	30	24	88	0	0
Dec-06	2	10	4	21	17	267	1	0
Jan-07	4	2	1	6	5	-40	0	0
Feb-07	0	2	1	6	5	200	1	0
Mar-07	0	0	0	0	0	0	0	0
Apr-07	1	0	0	0	0	-50	0	1
May-07	0	2	0	4	4	200	0	0
Jun-07	1	0	0	1	1	-50	0	0
Jul-07	0	2	0	4	4	200	0	0

Legend: L95%pib = lower 95% prediction interval boundary ; U95%pib = upper 95% prediction interval boundary ;%re = relative error of predictions expressed as percentage; w95%pi = width of 95% prediction interval; below = count of 1 if observation is lower than lower 95% prediction interval boundary; higher = count of 1 if observations higher than upper 95% prediction interval boundary

one month and two month horizons were made using the function "arima" in the package "stats" of the software "R", freely available from http://cran.r-project.org/ (accessed 2007). Initially, for all districts, a uniform SARIMA prediction model was used. This model had the parameters p=0, d=1, q=1, P=1, D=0, Q=1 (See Chapter 7 for model definition). This model was chosen based on preliminary testing on selected districts. SARIMA was performed on logarithmically transformed malaria case data after being augmented with the value one. Predictions and 95% prediction interval boundaries were back transformed, and rounded to positive integer values. Predictions were manually transferred to a text table compatible with "html" code for internet (http://www.iwmi.cgiar.org/health/malaria/predictions/) and E-mail posting, and posted within minutes, for RMOs to access.
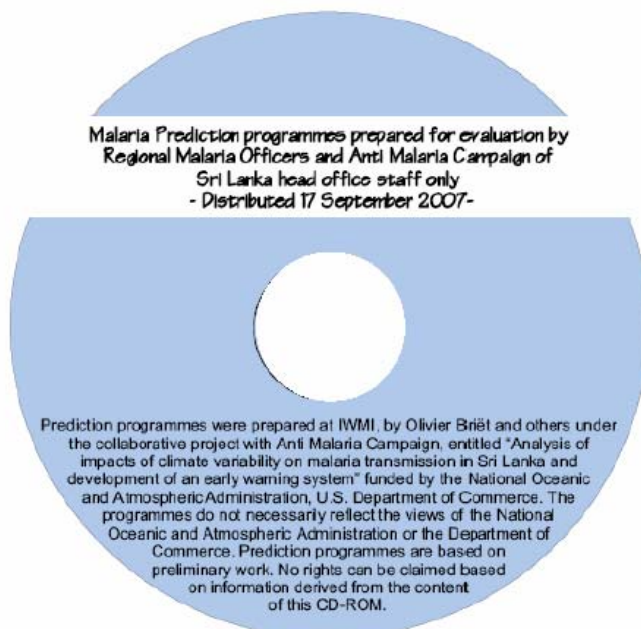
Figure 39 - CD-ROM

Cover image of CD-ROM provided to Anti Malaria Campaign head office staff and Regional Malaria Officers.

Forecasting system - Phase 2: Provision of software to RMOs and AMC staff for prediction

Present RMOs and AMC staff were provided with a CD-ROM (Figure 39) containing the software and data, and the AMC staff were given additional CD-ROMs for those RMOs unable to attend, for further distribution during a monthly review meeting.

The software could be easily installed following prompts on computers with a Windows based operating system. RMOs could enter their recent monthly malaria case data and append these to the data provided (monthly district case positives since January 1972 until July 2007). The CD-ROM contained a software program which makes predictions at a forecasting horizon of 1 to 12 months, including prediction intervals, using SARIMA models. By selecting their district (one operator switch), RMOs could generate predictions for their district within seconds. In addition, alternative recommended SARIMA model parameter settings were given (based on those found optimal in Chapter 7) for each district specifically. Apart from numerical output, also a graph is generated (Figure 40). In addition, the AMC head office was provided with a software macro program allowing calculation of predictions of all districts in a single action, making use of the recommended model parameter settings for each district. For the prediction program, the occurrence of missing data was allowed. In case the most recent data are missing for a district, the prediction for a future month is simply one step more "ahead".
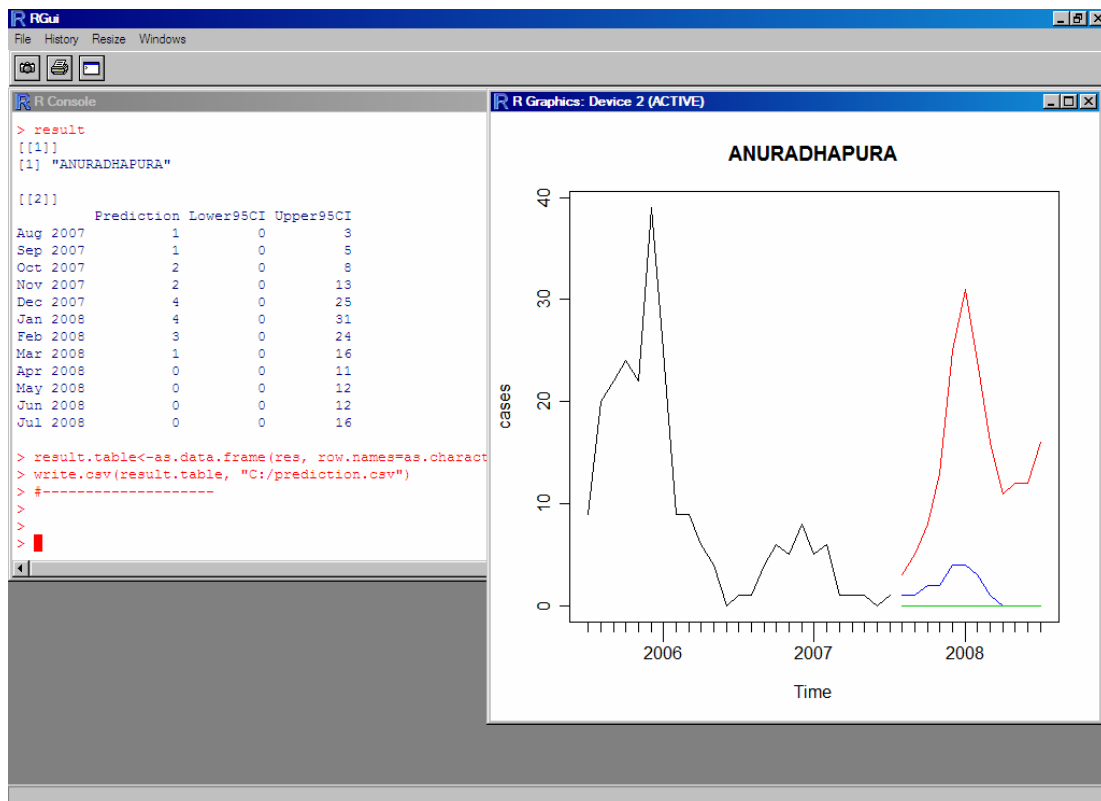
Figure 40 - Screen shot of output

Screen shot of an output of malaria case prediction program for a selected district.

Discussion

The models developed used case data only, as rainfall did not improve models substantially, and rainfall data were (still) cumbersome to process. Despite their simplicity, the models were received well by the AMC, as this was the first time such a system was tested in the country. It was felt that this was a first step towards developing useful models for malaria forecasting in Sri Lanka.

During the initial phase of the development of the model, it was noted that the success of a forecasting system would depend on the reporting of the cases data in time (by the end of each month) and having a good communication systems (Computers, internet and telephone facilities etc.) at the district level where operational activities are carried out. Such facilities were not equally distributed and the effects of the on-going civil war posed recurrent problems, especially in the north and east of the country. It was clear that the method of reporting case data quickly was of paramount importance if the models were to be used effectively. The system that operated within the AMC was that the district case data were shared at the monthly review meetings held during the last week of the month. It has been the experience that the full set of the data were not always gathered due to various problems that occurred at a district level. The district level counts were collected from different sub-centres, and the reporting system often depended on local transport or having an AMC vehicle for conveyance.

While the individual RMOs were very happy to use the tool in the field where possible (those who had facilities), it was realised that predictions were also necessary at the central planning level. In the initial phase it was also seen that the operationalization of such a system with predictions made at IWMI could not be sustained after the project period, therefore, it was agreed that the forecasting would be handled centrally, and results shared at review meetings, or communicated via phone if meetings were not held. This would be also beneficial in monitoring and sharing the information across districts, and keeping tabs on the outbreaks occurring near district boundaries.

Thus, the prediction system was handed over to the AMC at no additional cost for use after the end of the project. This was possible as the R software could be freely distributed under a GNU general public license. (Although the pre-made prediction

program macro is very user friendly, the software is otherwise not menu driven and therefore somewhat complicated). The macro program developed at IWMI was distributed as a global public good under IWMI's policy on intellectual property.

Table 13 - Observed values and predictions at a forecasting horizon of 1 month ahead for Trincomalee District

Month	Observed	Prediction	L95%pib	U95%pib	w95%pi	%re	below	higher
Aug-05	8	13	2	63	61	56	0	0
Sep-05	29	10	1	48	47	-63	0	0
Oct-05	23	35	7	162	155	50	0	0
Nov-05	24	24	5	112	107	0	0	0
Dec-05	11	26	5	121	116	125	0	0
Jan-06	8	12	2	57	55	44	0	0
Feb-06	10	10	1	47	46	0	0	0
Mar-06	14	12	2	58	56	-13	0	0
Apr-06	20	16	3	76	73	-19	0	0
May-06	12	22	4	102	98	77	0	0
Jun-06	2	13	2	61	59	367	0	0
Jul-06	2	2	0	15	15	0	0	0
Aug-06	1	3	0	17	17	100	0	0
Sep-06	3	2	0	11	11	-25	0	0
Oct-06	0	4	0	23	23	400	0	0
Nov-06	0	0	0	5	5	0	0	0
Dec-06	2	0	0	6	6	-67	0	0
Jan-07	4	3	0	18	18	-20	0	0
Feb-07	1	5	0	26	26	200	0	0
Mar-07	23	1	0	9	9	-92	0	1
Apr-07	22	27	5	129	124	22	0	0
May-07	15	20	4	93	89	31	0	0
Jun-07	6	14	2	67	65	114	0	0
Jul-07	4	6	1	31	30	40	0	0

Legend: as in Table 12

Some small outbreaks occurred that were not picked out by the predictive models. AMC head office staff noted that the models were not able to adequately forecast the beginning of small outbreaks such as that occurred in the Trincomalee district (Table 13), although the cases in ensuing months were adequately predicted. In this outbreak, the cases occurred primarily in migrant workers, and migration is a factor that was not directly accounted for in this model. With the low prevalence of malaria throughout Sri Lanka at the time of writing, malaria cases occurred erratically and were extremely difficult to forecast.

Within the framework of the model, the usefulness of the prediction system for actual control planning was difficult to estimate. The RMOs were of the view that the prediction intervals gave an indication of the control capacity that needed to be maintained to a level to be able to cope with possible case numbers in the higher end of the 95% prediction interval. Others expressed that developing the models to obtain predictions at a sub-district scale should be the next target, so that control operations could be carried out selectively. Many RMOs keep records at Grama Niladhari level, which is an administrative area level even below the MOH area level. A system incorporating sub district variability would be much more complex to install, especially as spatial dependence would have to be incorporated in order to cope with low and therefore erratic counts in time series. Furthermore, since there was a considerable delay in reporting of MOH area level data (much larger than the delay in reporting of district aggregated data), this would be of little use if predictions were made at the AMC head office.

Although (Bayesian) negative binomial models have now been developed (Chapter 8) for which data do not need to be transformed, and which may be more accurate and give more realistic prediction intervals than Gaussian models on logarithmically transformed data, these models need substantially more computing power than the (frequentist) Gaussian models to yield prediction estimates. Adequate computers that will yield results in a reasonable time period may not be available at the peripheral level. Also, experience with interpreting model convergence is often necessary, and such models are therefore not yet ready for use by staff with limited experience with Bayesian statistical software.

Chapter 8 – Generalized (S)ARIMA models for count data with application to malaria time series

This chapter is being prepared for submission to a biostatistics journal

Olivier JT Briët [*1,2], Penelope Vounatsou[2], Priyanie H Amerasinghe[3]

1 International Water Management Institute, P.O. Box 2075, Colombo, Sri Lanka, Telephone: +94112787404, Fax: +94112786854

2 Swiss Tropical Institute, Socinstrasse 57, P.O. Box CH-4002, Basel, Switzerland

3 International Water Management Institute Sub Regional Office for South Asia, c/o ICRISAT, Patancheru, AP 502 324, Andhra Pradesh, India

* Corresponding author

Abstract

Generalized autoregressive moving average (GARMA) models were extended to generalized multiplicative seasonal autoregressive integrated moving average (GSARIMA) models for the parsimonious and observation-driven modelling of non Gaussian, non stationary and / or seasonal time series data. Model fit was done by Bayesian Markov chain Monte Carlo methods. The choice of link function and methods for dealing with zeros amongst the observations were briefly examined. The model was demonstrated on monthly malaria episode count data in a district in Sri Lanka assuming that they follow a negative binomial distribution with a logarithmic link function. Computer code was provided for the implementation of these models in freely available software.

1 Introduction

Malaria has been a major public health problem in Sri Lanka [2] until incidence dwindled recently [135]. Malaria in Sri Lanka is seasonal and unstable and fluctuates in intensity both spatially and temporally, thus resources for control have to be spread in time and space to be prepared for outbreaks, which have occurred in the past despite very aggressive and effective malaria control operations [38]. Currently, the Anti Malaria Campaign Directorate of the Ministry of Health in Sri Lanka is testing a malaria forecasting system which could contribute to a more focussed approach for control, and could have a positive impact on the resource allocation for malaria control over space and time. This forecasting system uses multiplicative seasonal autoregressive integrated moving average (SARIMA) models which assume that logarithmically transformed monthly malaria case count data are approximately Gaussian distributed. Although Box-Cox class transformation of the malaria counts may yield approximately Gaussian distributed data, such approximation is less close for observations with a low expected mean [136]. Also, low count data may include zeros, which renders Box-Cox transformation not applicable. To overcome this problem, a small constant can be added to the data. Especially with low current monthly case counts, the Gaussian modelling on transformed data may result in inaccurate prediction intervals. Models assuming a negative binomial-class distribution for (untransformed) malaria count data may be more appropriate [127,137,138]. However, negative binomial-class models incorporating SARIMA structure were not yet available.

Benjamin and colleagues [139] provide a framework for generalized linear autoregressive moving average (GARMA) models, and they discuss amongst others GARMA(p,q) models for Poisson and negative binomially distributed data. GARMA models allow for lagged dependence in observations are also known as observation driven models. Alternatively, parameter driven models (also) allow dependence in latent variables [140-143]. The former models are easier to estimate and prediction is straightforward, while the latter models are in general easier to interpret [144,145]. Jung and colleagues [146] found that both types of models performed similarly.

GARMA models relate predictors and ARMA components to a transformation of the mean parameter of the data distribution (λ_t), via a link function. A log link function

ensures that λ_t is constrained to the domain of positive real numbers. Lagged observations should then also be logarithmically transformed, which is not possible for observations with value zero. Zeger and Qaqish [147] discuss adding a small constant to the data, either to all data, or only to zeros.

Grunwald [148] discuss a conditional linear autoregressive (CLAR) model with an identity link function. In order to ensure a positive λ_t, restrictions can be put on the parameters. In stationary models, the intercept can be restricted to be positive. Also, restricting the autoregressive coefficient ϕ_1 to $0 < \phi_1 < 1$ in a first order autoregressive Poisson model will ensure a positive λ_t. Brandt and Williams [149] discuss that negative values of ϕ_1 may be admissible depending on the size of the intercept and the variance of the series.

A variant of the GARMA models, generalized linear autoregressive moving average (GLARMA) models, is presented by Davis and colleagues [145].

Heinen [150] proposes a class of autoregressive conditional Poisson (ACP) models with methods allowing for over and under dispersion in the marginal distribution of the data. Another class of Poisson models with auto correlated error structure use "binomial thinning" and are called integer-valued autoregressive (INAR) models. The autoregressive parameter is restricted to $0 < \phi_1 < 1$. INAR models may be theoretically extended to moving average (INMA) and INARMA models [151,152], but these are not easily implemented [153].

An alternative parameter driven modelling approach is to assume an autoregressive process on time specific random effects introduced in the mean structure using a logarithmic link function [154]. Such a model is sometimes called a stochastic autoregressive mean (SAM) model [146] and has frequently been applied in Bayesian temporal and spatio-temporal modelling [49,50,138,144,155-157].

Of the models discussed above, the GARMA framework appears the most flexible for the modelling of count data with autoregressive and / or moving average structure. Benjamin and colleagues [139] apply a GARMA model to a time series of polio cases with a deterministic component for the seasonal effect using two sine and cosine pairs. However, if the seasonal component is thought to be stochastic the GARMA model presented by Benjamin and colleagues [139] is not appropriate. Also, many

time series of count data exist which are non stationary, and this is also the case for the malaria case count time series in districts in Sri Lanka. The GARMA model could easily be extended to a class of generalised multiplicative seasonal autoregressive integrated moving average (GSARIMA) models, analogous to SARIMA models for Gaussian distributed data. The class of GSARIMA models includes generalised autoregressive integrated moving average (GARIMA) models. GSARIMA models are defined in section 2 of this chapter. Model estimation through full Bayesian inference is briefly described. Some properties of the choice of link function and choices for data transformations in case of a logarithmic link function and observations with the value zero are investigated section 3. In section 4, a negative binomial GSARIMA model is applied to malaria count time series in a district in Sri Lanka. Section 5 provides concluding remarks and software code is provided in the Appendix.

2 The model

Model formulation

The negative binomial GARMA(p,q) model with logarithmic link function of Benjamin and colleagues [139] for a time series $\mathbf{y}^T = (y_t, y_{t+1}, ..., y_{t+n})$ of length n can be written:

$$y_t \sim NegBin(\lambda_t, \psi) \text{ where } f(y_t \mid \lambda_t, \psi) = \frac{\lambda_t^{y_t}}{y_t!} \frac{\Gamma(\psi + y_t)}{\Gamma(\psi)(\psi + \lambda_t)^{y_t}} \frac{1}{\left(1 + \frac{\lambda_t}{\psi}\right)^{\psi}} \quad (1)$$

$$\begin{aligned}
\log(\lambda_t) &= \Phi_p(B)\left[\mathbf{x}_t^T \boldsymbol{\beta} - \log(y'_t)\right] + \log(y'_t) \\
&\quad - \Theta_q(B)\log(y'_t / \lambda_t) + \log(y'_t / \lambda_t)
\end{aligned} \quad (2)$$

$$y'_{t-l} = \max(y_{t-l}, c), 0 < c \le 1 \quad (3)$$

where $\Phi_p(B) = 1 - \phi_1 B - ... - \phi_p B^p$, $\Theta_q(B) = 1 - \theta_1 B - ... - \theta_q B^q$, B is a backshift operator with $B^d(y_t) = y_{t-d}$, $\boldsymbol{\beta}^T = (\beta_0, \beta_1, \beta_2, ..., \beta_v)$ is a vector of coefficients for $\mathbf{x}_t^T = (x_0, x_{1,t}, x_{2,t}, ..., x_{v,t})$, which is a vector of an intercept multiplier (usually taken as $x_0 = 1$) and v time dependent covariates. Under the specification for the negative binomial distribution in Eq 1, $E(y_t) = \lambda_t$ and $V(y_t) = \lambda_t + \frac{\lambda_t^2}{\psi}$. As $\psi \to \infty$, the limiting form of the negative binomial distribution is the Poisson distribution. The

transformation in Eq 3, henceforth called "ZQ1" is proposed by Zeger and Qaqish [147] in order to avoid problems of taking the logarithm of observations with value zero. Zeger and Qaqish [147] also suggest an alternative method, henceforth called "ZQ2", which translates into the model variant:

$$\log(\lambda_t) = \Phi_p(B)\left\{\log\left[\exp\left(\mathbf{x}_t^T\boldsymbol{\beta}\right)+c\right]-\log(y_t+c)\right\}+\log(y_t+c)$$
$$-\Theta_q(B)\log\left[(y_t+c)/(\lambda_t+c)\right]+\log\left[(y_t+c)/(\lambda_t+c)\right]$$

The above models can be extended to GSARIMA(p,d,q,P,D,Q) models including seasonality and integration. With ZQ1 transformation, the extended model is:

$$\log(\lambda_t) = \Phi_p(B)(1-B)^d\left(1-B^s\right)^D\Phi_P^*\left(B^s\right)\left[\mathbf{x}_t^T\boldsymbol{\beta}-\log(y'_t)\right]+\log(y'_t)$$
$$-\Theta_q(B)\Theta_Q^*\left(B^s\right)\log(y'_t/\lambda_t)+\log(y'_t/\lambda_t)$$

where s is the length of the period ($s=12$ for monthly data with an annual cycle), $\Phi_P^*\left(B^s\right)=1-\phi_1^*B^s-...-\phi_P^*B^{sP}$, $\Theta_Q^*\left(B^s\right)=1-\theta_1^*B^s-...-\theta_Q^*B^{sQ}$, $\Phi_p(B)$, $\Theta_q(B)$, and B are as above. Similarly, with ZQ2 transformation, the extended model is:

$$\log(\lambda_t) = \Phi_p(B)(1-B)^d\left(1-B^s\right)^D\Phi_P^*\left(B^s\right)\left\{\log\left[\exp\left(\mathbf{x}_t^T\boldsymbol{\beta}\right)+c\right]-\log(y_t+c)\right\}+\log(y_t+c)$$
$$-\Theta_q(B)\Theta_Q^*\left(B^s\right)\log\left[(y_t+c)/(\lambda_t+c)\right]+\log\left[(y_t+c)/(\lambda_t+c)\right]$$

In the GARMA framework, negative binomially distributed data could be modelled via a logarithmic link function or an identity link function, whichever is thought most appropriate for the series. A negative binomial GSARIMA(p,d,q,P,D,Q) model with identity link function takes the form:

$$\lambda_t = \Phi_p(B)(1-B)^d\left(1-B^s\right)^D\Phi_P^*\left(B^s\right)\left[\exp\left(\mathbf{x}_t^T\boldsymbol{\beta}\right)-y_t\right]+y_t$$
$$-\Theta_q(B)\Theta_Q^*\left(B^s\right)(y_t-\lambda_t)+(y_t-\lambda_t)$$
.

Model estimation

Benjamin and colleagues [139] employ maximum likelihood estimation through iterative weighted least squares and base inference on asymptotic results. In order to overcome computational problems associated with the maximum likelihood approach, we have formulated the model in a Bayesian framework. In a fully Bayesian context, the joint posterior distribution conditional on the first w observations with w equal to the maximum autoregressive lag required in the model is proportional to the product of the likelihood and the prior distributions of the parameters, which are a priori

considered independent:

$$p\left(\psi,\boldsymbol{\beta},\boldsymbol{\varphi},\boldsymbol{\theta},\boldsymbol{\varphi}^*,\boldsymbol{\theta}^* \mid y_1,...,y_n,\mathbf{X}\right)\infty$$

$$p\left(y_{w+1},...,y_n \mid y_1,...,y_w,\mathbf{X},\psi,\boldsymbol{\beta},\boldsymbol{\varphi},\boldsymbol{\theta},\boldsymbol{\varphi}^*,\boldsymbol{\theta}^*\right)p\left(\psi\right)p\left(\boldsymbol{\beta}\right)p\left(\boldsymbol{\varphi}\right)p\left(\boldsymbol{\theta}\right)p\left(\boldsymbol{\varphi}^*\right)p\left(\boldsymbol{\theta}^*\right)$$

with $\mathbf{X}$ an n by $v+1$ matrix of covariates, $\boldsymbol{\varphi}^T=\left(\phi_1,...,\phi_p\right)$, $\boldsymbol{\theta}^T=\left(\theta_1,...,\theta_q\right)$,

$\boldsymbol{\varphi}^{*T}=\left(\phi_1^*,...,\phi_P^*\right)$, and $\boldsymbol{\theta}^{*T}=\left(\theta_1^*,...,\theta_Q^*\right)$. For the purpose of inference, stationary and

invertibility conditions discussed by Box and Jenkins [158] were considered.

In Bayesian inference, prior distributions need to be assigned to all model parameters.
We assume a stationary model and therefore the auto correlation and moving average
parameters were given prior distributions that match the stationarity and invertibility
region, which were constructed using an algorithm provided by Jones [159]. For this
purpose, we reparameterize the likelihood in terms of $\mathbf{r}$, $\mathbf{r}^T=\left(r_1,...,r_p\right)$ instead of $\boldsymbol{\varphi}$,

where $\phi_p=2r_p-1$ and $\phi_{p-i}=2r_{p-i}-1-\sum_{k=1}^{i}\left(2r_{p-k}-1\right)\left(2r_{p-k+1}-1\right),i=1,...,p-1$ and

assume the following prior distributions: $r_i \sim Beta\left(\left[\tfrac{1}{2}(i+1)\right],\left[\tfrac{1}{2}i+1\right]\right),i=1,...,p$,

where $[x]$ denotes the integer part of x. Parameters $\boldsymbol{\theta}$, $\boldsymbol{\varphi}^*$, and $\boldsymbol{\theta}^*$ are also

reparameterized and similar prior distributions discussed above are assumed on the
new parameterization. Further priors chosen were $\beta_0, ..., \beta_v \sim U(-\infty, \infty)$ and $\psi \sim$
Ga(0.01, 0.01). For the first w observations, the residuals on the predictor scale (e.g.
$\log(y'_t) - \log(\lambda_t)$ in the case of a logarithmic link function) were set to zero. In a
Bayesian framework, for use in the identity link function, a restriction can be put on
the mean λ_t itself, that is $\lambda_t \geq 0$. In this study the GSARIMA models were estimated
using the free Bayesian software "WinBUGS"[55], which uses Gibbs sampling.
Congdon [160] shows some examples of analysis of non-Gaussian time series using
WinBUGS. It should be noted that WinBUGS simulates each node in turn, which can
make convergence very slow and the program very inefficient for some time series
models. Model building in WinBUGS was straightforward although slightly
laborious for complex SARIMA structures. An example of WinBUGS code (written
for use with the R package "R2WinBUGS"[161]) for a negative binomial GSARIMA
model with logarithmic link function and ZQ1 transformation is given in the
Appendix).

3 Simulation studies

Three simulation studies were performed. In the first study, we assessed the influence of the choice of link function and choices for data transformations on the distribution of simulated data. In the second study, the effect of (mis)specification of the link function and data transformation when estimating GARMA model parameters was investigated. In the third study, the ability to estimate simulated data series with GSARIMA structure was briefly explored. Data series with (invertible) GSARIMA structure were simulated in the software environment R by writing the GSARIMA model in the form of an infinite AR representation, approximated by a finite order AR representation [162].

The choice of link function and choices for data transformations in case of a logarithmic link function and observations with the value zero

For Poisson AR(1) models with identity link function and logarithmic link function (with transformation ZQ1 and ZQ2), the distribution properties of simulated series of length 1 000 000 were compared for an intercept $\exp(\beta_0) = 2$, $\exp(\beta_0) = 10$ or $\exp(\beta_0) = 100$, and a coefficient $\phi_1 = -0.5$ or $\phi_1 = 0.5$, and a constant $c = 0.1$, or $c = 1$ (Table 14).

Table 14 - Distribution properties of simulated series of different Poisson AR(1) models

Model				mean	Variance	skewness	kurtosis
Link	Exp(β_0)	ϕ_1	c				
identity	2	0.5		2.00	2.67	1.06	4.54
log-ZQ1	2	0.5	0.1	1.54	2.32	1.14	4.41
log-ZQ1	2	0.5	1.0	1.98	2.35	0.91	4.02
log-ZQ2	2	0.5	0.1	1.55	2.30	1.12	4.29
log-ZQ2	2	0.5	1.0	1.91	2.15	0.84	3.80
identity	2	-0.5		2.01	2.62	0.83	3.63
log-ZQ1	2	-0.5	0.1	3.01	10.05	1.73	5.90
log-ZQ1	2	-0.5	1.0	2.14	2.50	0.82	3.73
log-ZQ2	2	-0.5	0.1	3.03	10.47	1.75	5.94
log-ZQ2	2	-0.5	1.0	2.16	2.55	0.87	3.95
identity	10	0.5		10.00	13.37	0.47	3.31
log-ZQ1	10	0.5	0.1	9.61	13.37	0.40	3.18
log-ZQ1	10	0.5	1.0	9.62	13.34	0.40	3.17
log-ZQ2	10	0.5	0.1	9.63	13.23	0.39	3.17
log-ZQ2	10	0.5	1.0	9.75	12.50	0.40	3.16
identity	10	-0.5		10.00	13.34	0.37	3.14
log-ZQ1	10	-0.5	0.1	10.42	17.30	2.40	41.83
log-ZQ1	10	-0.5	1.0	10.41	16.00	0.84	4.99
log-ZQ2	10	-0.5	0.1	10.41	16.80	2.14	36.58
log-ZQ2	10	-0.5	1.0	10.31	13.99	0.61	3.97
identity	100	0.5		99.98	133.58	0.15	3.04
log-ZQ1	100	0.5	0.1	99.67	133.47	0.12	3.02
log-ZQ1	100	0.5	1.0	99.68	133.23	0.12	3.01
log-ZQ2	100	0.5	0.1	99.66	133.21	0.13	3.02
log-ZQ2	100	0.5	1.0	99.66	132.45	0.13	3.02
identity	100	-0.5		100.00	133.27	0.11	3.01
log-ZQ1	100	-0.5	0.1	100.34	134.70	0.18	3.05
log-ZQ1	100	-0.5	1.0	100.34	134.72	0.18	3.06
log-ZQ2	100	-0.5	0.1	100.34	134.68	0.18	3.07
log-ZQ2	100	-0.5	1.0	100.33	134.13	0.18	3.06

Legend: ZQ1: transformation method corresponding to equation 2.2 in Zeger and Qaqish [147]; ZQ2: transformation method corresponding to equation 2.4 in Zeger and Qaqish [147].

For a (high) intercept of $\exp(\beta_0) = 100$, all models had near identical results, although the log-link models resulted in slightly lower mean, variance, skewness and kurtosis with a positive ϕ_1, and vice versa for a negative ϕ_1. This effect became stronger at a lower intercept of $\exp(\beta_0) = 10$. The impact of the choice for constant c (in the log-link models) was strong at the lower intercepts $\exp(\beta_0) = 2$ and $\exp(\beta_0) = 10$ (for the latter, for $\phi_1 = -0.5$), and a value of $c = 1$ gave results most similar to the model with identity link function. The choice for the transformation method gave variable results: at a low intercept $\exp(\beta_0) = 2$ (and $c = 1$), the mean, variance, skewness and kurtosis were more similar to the identity link model for ZQ1 (except the variance for $\phi_1 = -0.5$), whereas at the intercept of $\exp(\beta_0) = 10$ (and $c = 1$), the mean, variance, skewness and kurtosis were more similar to the identity link model for ZQ2 (except the variance and kurtosis for $\phi_1 = 0.5$).

Effect of (mis)specification of the link function

The effect of choice of the link function, of the ZQ transformation and of the value of the parameter c on parameter estimates was studied on a simulated Poisson time series of length 1 000, with AR($p = 1$) structure with $\phi_1 = 0.5$, with a logarithmic link function using transformation method ZQ1 with $c = 1$, and an intercept $\exp(\beta_0) = 2$. Models were estimated using three chains with each a length of 2 000 iterations, including a burn-in of 1 000 iterations (Table 15).

Table 15 – Parameter estimates and 95% credible intervals for three types of models on a simulated Poisson AR(1) series with log link function, "ZQ1" transformation, intercept = 2, $c = 1$, and $\phi_1 = 0.5$.

Link	c	intercept	ϕ_1	*maref*	DIC
identity	NA	1.98 (1.86 – 2.10)	0.33 (0.26 – 0.39)	0.48	3360
log-ZQ1	0.1	2.13 (2.01 – 2.25)	0.18 (0.14 – 0.22)	0.52	3402
log-ZQ1	1	2.00 (1.85 – 2.17)	0.45 (0.38 – 0.53)	0.46	3346
log-ZQ2	0.1	2.12 (2.01 – 2.25)	0.17 (0.13 – 0.22)	0.52	3406
log-ZQ2	1	2.07 (1.95 – 2.21)	0.46 (0.37 – 0.55)	0.50	3371

The model "log-ZQ1" with $c = 1$ performs best, as expected. The identity link model appears to do better than the model "log-ZQ2" with $c = 1$, based on the DIC and *maref*, but for the identity link model, the 95% credible interval of ϕ_1 was below 0.5, which was the value used for simulation. With $c = 0.1$, for both ZQ1 and ZQ2

transformations, the 95% credible intervals for both the intercept and ϕ_1 did not include the parameter values used for the simulation.

Ability to estimate GSARIMA structure

Finally, a negative binomial GSARIMA(2,1,0,0,0,1) time series of length 1 000 was simulated, with a logarithmic link function, ZQ1 transformation with $c = 1$, an external variable sampled $x_t \sim N(0,1)$, and $\beta_1 = 0.7$, $\phi_1 = 0.5$, $\phi_2 = 0.2$, $\theta_1^* = 0.5$, $s = 12$, $\psi = 5$. This model was estimated with a chain of a length of 600 iterations including a burn-in of 200 iterations. The results in Table 16 show that the method proposed here was able to estimate the parameters correctly.

Table 16 – Parameter estimates and 95% credible intervals on a simulated negative binomial GSARIMA(2,1,0,0,0,1) time series with log link function, zq1 zero transformation, c parameter 1, $\beta_1 = 0.7$, $\phi_1 = 0.5$, $\phi_2 = 0.2$, $\theta_1^* = 0.5$, $s = 12$, and $\psi = 5$.

Parameter	Estimate
β_1	0.71 (0.69 – 0.72)
ϕ_1	0.45 (0.39 – 0.51)
ϕ_2	0.23 (0.16 – 0.29)
θ_1^*	0.48 (0.44 – 0.52)
ψ	5.05 (4.46 – 5.63)

4 Application to malaria time series analysis

This section provides an example of a GSARIMA model applied to monthly malaria case count y_t for the period 1972 – 2005 in the district of Gampaha in Sri Lanka (Figure 41), with rainfall as covariate. Records of malaria positive blood films were reported monthly by government health facilities and aggregated by the Anti Malaria Campaign (AMC) of Sri Lanka. Rainfall was the monthly district average height of the precipitation column, which was derived from monthly island-wide precipitation surfaces. These rainfall surfaces were generated by spatial interpolation of precipitation records collected by 342 stations across the island. The data was earlier described by Briët and colleagues [111], who fitted Gaussian SARIMA and ARIMA models with a deterministic seasonal effect to logarithmically transformed malaria case data.

Preliminary model identification using frequentist Gaussian SARIMA

Because Bayesian model fit using Markov chain Monte Carlo (MCMC) algorithms is computationally expensive, rather than fitting many possible MCMC models, preliminary model identification was performed using standard frequentist tools developed for time series with Gaussian marginal errors. A visual analysis of the malaria time series (Figure 41) detected the presence of a long term (inter annual) change in the mean level, unstable variance which appears to increase with the mean, and multiplicative seasonality (the size of the seasonal effect is proportional to the mean). For the purpose of choosing the parameters *p, d, q, P, D*, and *Q*, the data was transformed using a fitted Box-Cox transformation in order to stabilize the variance, to make the seasonal effect additive, and to make the data approximately normally distributed. The fitted Box-Cox parameters were a power of 0.24 and a constant of 0.25 added to each observation prior to transformation. In the transformed time series the presence of long term change in the mean level was also apparent. Although the changes in the mean level could potentially be related to malaria control efforts, development of parasite and vector resistance, etc., such data were not considered here. Hence the trend was treated as a stochastic trend, which was (first order) difference stationary. The augmented Dickey – Fuller test supported the presence of a unit-root (p = 0.14). Plots of the auto correlation function (ACF) (not shown) and the partial auto correlation function (PACF) (not shown) of the differenced series showed significant (partial) auto correlation at lags of three and twelve months. Gaussian SARIMA models and ARIMA models with a deterministic component for the seasonal effect with *d* = 1 were fitted with the (frequentist) software R. The covariate matrix for the deterministic component for the seasonal effect using second order harmonics (*i.e.* using two sine and cosine pairs) is given by

$\mathbf{x}_t^T = \left[\sin\left(2\pi t/12\right), \cos\left(2\pi t/12\right), \sin\left(2\pi t/6\right), \cos\left(2\pi t/6\right) \right]$. A (time independent) intercept was not included because the intercept drops out of the equation after first order differencing. Based on the preliminary analysis, four Gaussian SARIMA models and two Gaussian ARIMA models with second order harmonics were initially selected, based on Akaike's information criterion (Table 17).

Table 17 - Akaike's information criterion (AIC) for selected (Gaussian) models on Box-Cox transformed data. For all these models, where applicable, the autoregressive (ϕ_1 and ϕ_2) or moving average parameters (θ_1 and θ_2) corresponding to the first two lags were omitted.

Model	Excluding rainfall AIC	Including rainfall AIC
SARIMA(3,1,0,1,0,0)	1572.75	1574.55
SARIMA(3,1,0,0,0,1)	1573.09	1574.93
SARIMA(0,1,3,1,0,0)	1572.50	1574.45
SARIMA(0,1,3,0,0,1)	1572.84	1574.81
ARIMA(3,1,0) SOH	1566.17	1564.91
ARIMA(0,1,3) SOH	1565.10	1564.20

Legend: SOH: second order harmonics
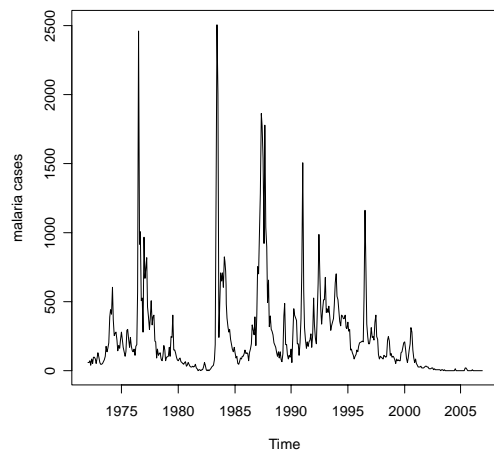
Figure 41 - Malaria cases in Gampaha District over time

Model identification using GSARIMA

Bayesian negative binomial versions of the first three SARIMA models and two ARIMA models with second order harmonics identified in the preliminary analysis were implemented in WinBUGS on untransformed data. The fourth SARIMA selected model was not implemented because for the estimation of more than one moving average parameter, in WinBUGS, the computing time per iteration increases exponentially with an increasing number of observations and are impractical with currently available computing power. Models were evaluated on two criteria. The first criterion was the deviance information criterion (DIC) which was calculated as the mean of the posterior distribution of the deviance conditional on the first ω observations (*D*), with ω equal to the maximum *w* of the five models compared, augmented with the number of effective estimated parameters as penalty to prevent over fitting. Models with lower DIC are considered better. The model with the

largest lag required ($w = 16$) was the model GSARIMA(3,1,0,1,0,0) with $s = 12$. A second criterion which was defined as the mean absolute relative error of fitted values

(*maref*): $maref = \sum_{t=\omega+1}^{t=n} \left| \frac{y_t - \hat{y}_t}{y_t + 1} \right| / (n - \omega)$ where $\hat{y}_t$ is the fitted number of malaria

positive cases at time t. Since the posterior distributions of the fitted values were skewed, the median of the posterior distribution of the fit was taken for $\hat{y}_t$. The *maref* is similar to the mean absolute percentage error, which is applicable for series for which the variance is dependent on the mean [90]. The *maref* statistic does not have a built in penalty to prevent over fitting, but among models with similar value of *maref*, the model with the least number of parameters should be preferred. The *maref* does not depend on distribution assumptions, in contrast to the DIC. Models were run with a single Markov chain of 11 000 iterations including a burn-in of 1 000 iterations, using WinBUGS default settings. Convergence was assessed in separate runs of 2 000 iterations with three chains, by studying plots of the Gelman-Rubin convergence statistic (on estimated parameters) as modified by Brooks and Gelman [54].

Based on the DIC, the best model was the negative binomial GARIMA(3,1,0) model with parameters for the first two lags (ϕ_1 and ϕ_2) omitted (fixed to zero), with deterministic harmonic seasonality and with rainfall preceding malaria with two months (Table 18). The parameter and deviance estimates for this model, henceforth "GARIMA(3,1,0)-SOH-RF", are detailed in Table 19. However, based on the *maref*, the negative binomial GSARIMA(3,1,0,1,0,0) model, also with the parameters for the first two lags omitted but without rainfall, was preferred. This model is also detailed in Table 19. Despite the GSARIMA(3,1,0,1,0,0) model having a higher (worse) DIC than the GARIMA(3,1,0)-SOH-RF model, the *maref* of the GSARIMA(3,1,0,1,0,0) model was one percent better than the *maref* of the GARIMA(3,1,0)-SOH-RF model, and required less than half the number of fitted parameters.

Table 18 - Selection criteria statistics for selected negative binomial models with transformation method "ZQ1" for logarithmic link models with $c = 1$. For all models, where applicable, the autoregressive (ϕ_1 and ϕ_2) or moving average parameters (θ_1 and θ_2) corresponding to the first two lags were omitted.

Model	nep	DIC	*maref*
GSARIMA(3,1,0,1,0,0)	3	4350.69	0.3876
GSARIMA(3,1,0,0,0,1)	3	4351.11	0.3889
GSARIMA(0,1,3,1,0,0)	3	4352.31	0.3881
GSARIMA(3,1,0,1,0,0)-RF	4	4352.48	0.3895
GARIMA(3,1,0)-SOH	6	4335.59	0.3961
GARIMA(0,1,3)-SOH	6	4336.10	0.3938
GARIMA(3,1,0)-SOH-RF	7	4333.29	0.3884
GARIMA(0,1,3)-SOH-RF	7	4334.11	0.3876

Legend: nep; number of estimated parameters; DIC: Deviance Information Criterion; *maref*: mean absolute relative error of fitted values; RF: with rainfall lagged at 2 months; SOH: second order harmonics.

Table 19 - Parameter estimates (mean and 95% credible interval) of selected negative binomial models

Parameter	GARIMA(3,1,0)-SOH-RF	GSARIMA(0,1,3,1,0,0)
β_{rain}	-0.335 (-0.656, -0.024)	
$\beta_{sin(2\pi t/12)}$	-0.103 (-0.233, 0.024)	
$\beta_{cos(2\pi t/12)}$	-0.148 (-0.278, -0.017)	
$\beta_{sin(2\pi t/6)}$	0.136 (0.058, 0.214)	
$\beta_{cos(2\pi t/6)}$	0.156 (0.073, 0.237)	
φ_3	-0.097 (-0.194, -0.001)	-0.133 (-0.227, -0.037)
$\varphi_1{}_*$		0.121 (0.027, 0.216)
ψ	4.542 (3.867, 5.267)	4.342 (3.692, 5.070)
Amplitude AH$^\$$	0.193 (0.07, 0.317)	
Amplitude SAH$^\$$	0.210 (0.127, 0.294)	
Phase shift AH$^\$$	4.829 (3.295, 6.345)	
Phase shift SAH$^\$$	-0.687 (-1.053, -0.335)	

Legend: GARIMA(3,1,0)-SOH-RF = GARIMA(3,1,0) model with parameters for the first two lags (ϕ_1 and ϕ_2) omitted, second order harmonics and rainfall lagged at 2 months (in m);

GSARIMA(3,1,0,1,0,0) = GSARIMA(3,1,0,1,0,0) model with parameters for the first two lags (ϕ_1 and ϕ_2) omitted; AH = annual harmonic, SAH = semi-annual harmonic; $^\$$ = derived parameter, phase shift = phase shift of the cosine function expressed in months.

Residual analysis

After fitting the GARIMA(3,1,0)-SOH-RF and GSARIMA(3,1,0) models, for each observation, the cumulative probability that the observed value occurred given its posterior distribution was calculated. Because of the fact that the posterior cumulative distribution function is discrete, the cumulative probability value was randomized by drawing a random value p_t from the uniform distribution in the interval $\left[F\left(y_t - 1, \lambda_t, r\right), F\left(y_t, \lambda_t, r\right) \right]$, where $F\left(y_t, \lambda_t, r\right)$ is the cumulative posterior distribution function (estimated with 10 000 samples from this distribution), following

a procedure by Dunn and Smyth [163]. This procedure is advocated by Benjamin and colleagues [139] for discrete GARMA models.

Figure 42 shows that for randomized cumulative probability values below about 0.4 for the GARIMA(3,1,0)-SOH-RF model and 0.5 for the GSARIMA(3,1,0) model, on average, cumulatively fewer observations had these values than expected based on the posterior density distributions. Therefore, on average, the part of the posterior density distributions below the median was too much spread out to the left. The lower boundaries of credibility intervals of the distributions were thus on average too low. For both models, probability values between about 0.5 and 0.9, the cumulative distribution function followed the diagonal. However, above a value of about 0.9, again the cumulative distribution function was (slightly) below the diagonal, indicating that the upper boundaries of credibility intervals were on average (slightly) too low.

The randomized cumulative probability values were converted into normalized randomized quantile residuals ε_t using the quantile function (inverse cumulative distribution function) of the normal distribution with zero mean and unity variance. Prior to conversion, randomized cumulative probability values of zero (when all 10 000 samples from the posterior distribution function were above the observed value) were set to 0.00 001 and randomized cumulative probability values of one (when all 10 000 samples from the posterior distribution function were below the observed value) were set to 0.99 999.
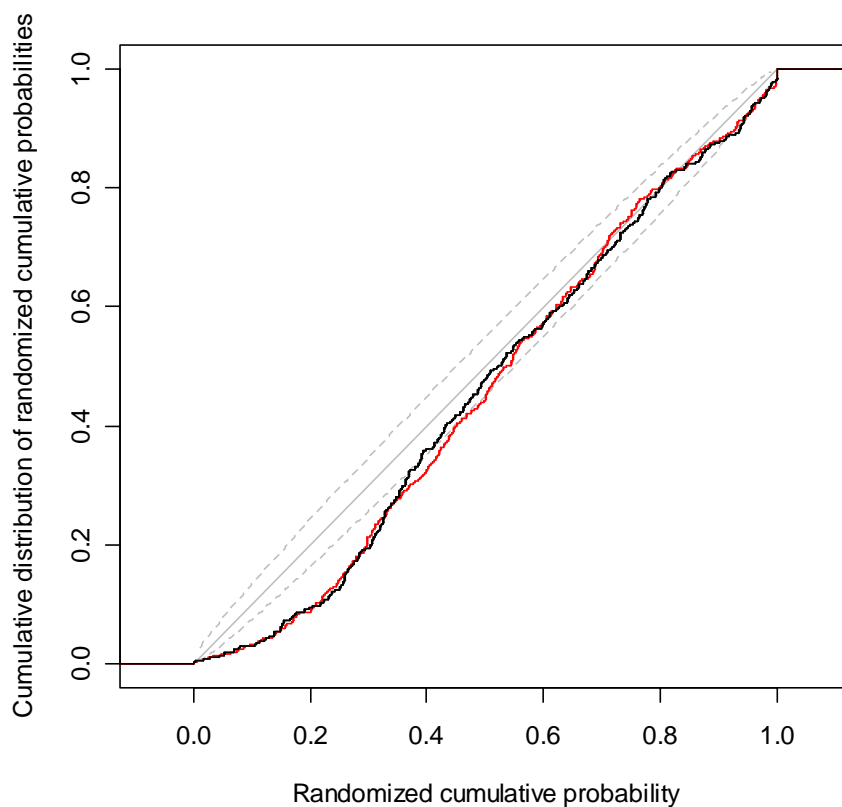
Figure 42 - Cumulative distribution function of randomized residual probabilities

Legend: The black line represents the cumulative distribution function of randomized residual probabilities of the GARIMA(3,1,0)-SOH-RF model. The red line represents the cumulative distribution function of randomized residual probabilities of the selected GSARIMA(3,1,0) model. The grey diagonal line (cumulative distribution equals randomized probability) represents on average appropriate posterior distributions. Dotted lines represent 95% confidence boundaries for proportions equalling the probability for 392 observations.

Figure 43 shows the normal Q-Q plots for the normalized randomized quantile residuals of the GARIMA(3,1,0)-SOH-RF model, for which the distribution is slightly leptokurtotic. A plot of these normalized randomized quantile residuals against time (Figure 44) appears a random scatter at first sight, but at closer inspection, extreme residuals occur more often during periods with more strong relative changes.
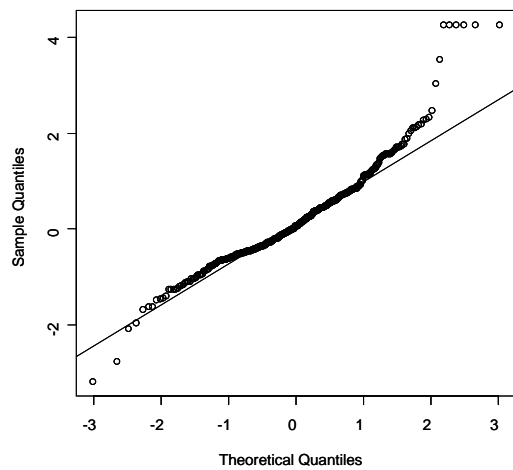
Figure 43 - Normal Q-Q plot of normalized randomised quantile residuals of the selected GARIMA(3,1,0)-SOH-RF model
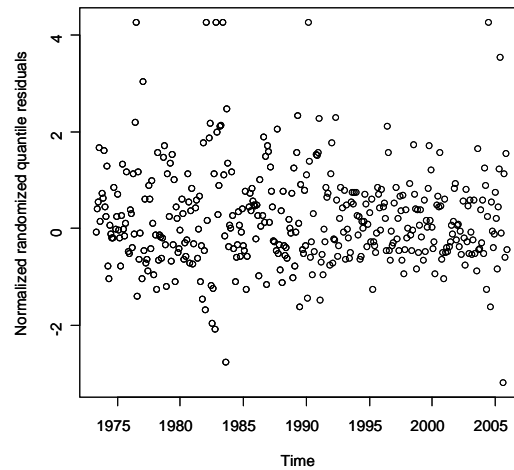
Figure 44 - Plot of normalized randomised quantile residuals of the selected GARIMA(3,1,0)-SOH-RF model against time

This is because the residuals ε_t are positively correlated with a relative change in malaria cases, with linear regression line $\varepsilon_t = 1.78 \log \left(\dfrac{y_t}{y_{t-1}} \right) + 0.22$ (Figure 45). The fact that this line does not go through the origin but has a (small but significant; p < 0.05) positive intercept is another indication that the posterior distributions have on average too much mass to the left, and therefore on average overestimate the residuals. Plots of the residuals of the GSARIMA(3,1,0) model looked very similar and are not shown.

First order differenced logarithmically transformed malaria cases
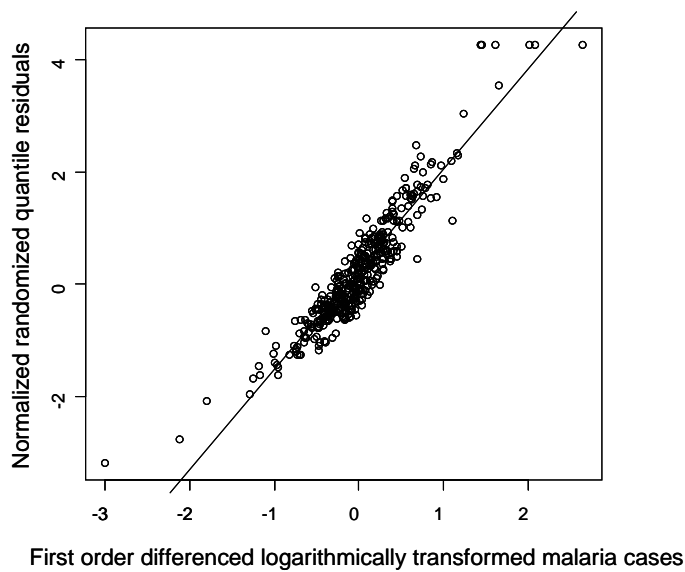
Figure 45 - Plot of normalized randomised quantile residuals of the selected GARIMA(3,1,0)-SOH-RF model against the logarithm of relative change

Figure 46 and Figure 47 show plots of the autocorrelation and partial autocorrelation functions of the normalized randomized quantile residuals of the GARIMA(3,1,0)-SOH-RF model. Although at lag 23 there seemed to be significant (partial) autocorrelation, the Ljung-Box test [164] showed no evidence of residual autocorrelation. The Ljung-Box statistic was 24.93 based on 24 lags which was not significant (p = 0.35) because the quantile corresponding to the 95th percentile of a chi-squared distribution with 23 degrees freedom (24 degrees minus one fitted ARMA parameter) is 35.17. The Ljung-Box test is valid under these mild conditions of non-normality, although for stronger non-normality, the Ljung-Box test is not robust and tends to reject the null hypothesis of no autocorrelation too quickly [165].
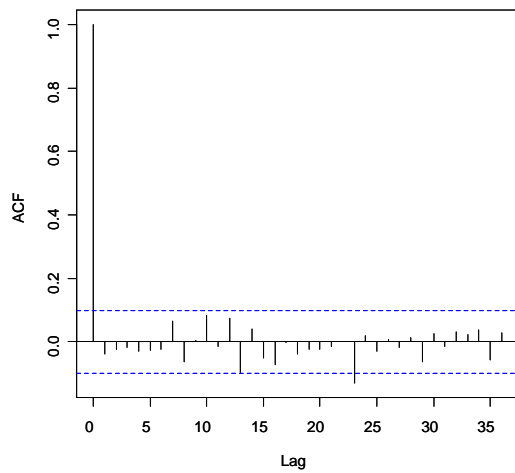
Figure 46 - Plot of the autocorrelation function of normalized randomised quantile residuals of the selected GARIMA(3,1,0)-SOH-RF model
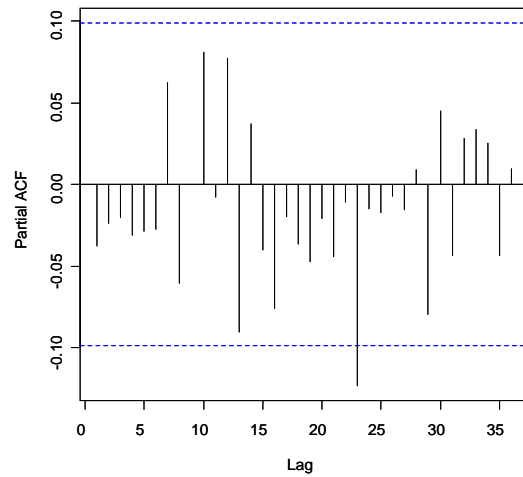
Figure 47 - Plot of the partial autocorrelation function of normalized randomised quantile residuals of the selected GARIMA(3,1,0)-SOH-RF model
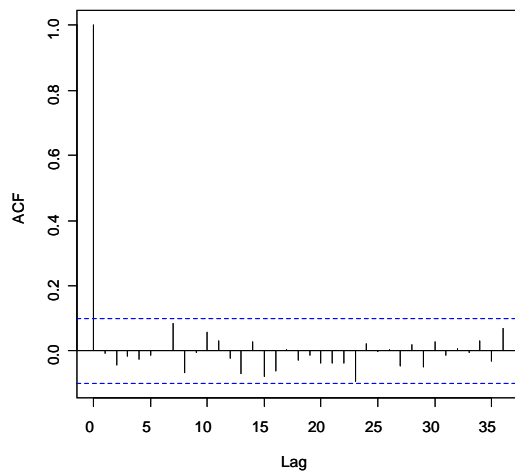
Figure 48 - Plot of the autocorrelation function of normalized randomised quantile residuals of the selected GSARIMA(3,1,0) model
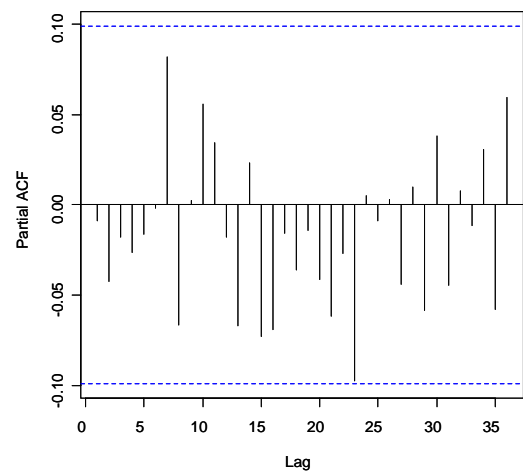
Figure 49 - Plot of the partial autocorrelation function of normalized randomised quantile residuals of the selected GSARIMA(3,1,0) model

Figure 48 and Figure 49 plot the autocorrelation and partial autocorrelation functions of the normalized randomized quantile residuals of the GSARIMA(3,1,0) model. For the residuals of this model, none of the lags showed significant (partial) autocorrelation, and the Ljung-Box statistic was 20.19 based on 24 lags was also not significant (p = 0.57) because the quantile corresponding to the 95th percentile of a

chi-squared distribution with 22 degrees freedom (24 degrees minus two fitted ARMA parameters) is 33.92. Based on the visual comparison of plots of the ACF and PACF, and on the size of the Ljung-Box statistic, the GSARIMA(3,1,0) model was better at modelling the autocorrelation present in this malaria time series.

For the purpose of (one step ahead) malaria prediction in practice, the simpler GSARIMA(3,1,0) model which does not require rainfall data collection may be preferable. Although the posterior distributions of the GSARIMA(3,1,0) model were slightly less appropriate than those of the GARIMA(3,1,0)-SOH-RF model, the 95% credibility interval boundaries are similar (Figure 42). The fact that the cumulative distribution functions do not perfectly match the diagonal in Figure 42 indicates that there is still room for improvement, both through modelling of a more complex autocorrelation structure (e.g. through time varying SARIMA parameters) and through inclusion of covariates. It is also possible that the assumption of an under-laying negative binomial distribution is not entirely appropriate. In the latter case, the DIC, which was based on this assumption, has less value than the *maref*. Apart from the fact that the *maref* does not depend on the assumption of a true under laying distribution, it is easier to interpret by malaria control staff.

5 Conclusions

For the modelling of monthly time series of counts of new malaria episodes in a district in Sri Lanka, GSARIMA models and GARIMA model with a deterministic seasonality component were developed. GSARIMA models and GARIMA models are an extension of a class of GARMA models, [139], and are suitable for parsimonious modelling of non-stationary seasonal time series of (over dispersed) count data with negative binomial conditional distribution.

Models were presented with a choice of identity link function or logarithmic link function. For the identity link function, an alternative to commonly used parameter restrictions was proposed. For models with a logarithmic link function, the choice of the transformation method for dealing with observations with a value of zero and its threshold parameter was studied. With a high intercept of $\exp(\beta_0) = 100$, simulated data series with an AR($p = 1$) model had near identical distributions. However, with lower intercepts of $\exp(\beta_0) = 10$ and $\exp(\beta_0) = 2$, the distributions varied in shape. Especially with a negative value of ϕ_1, at low intercepts, a value of $c = 0.1$ caused the

distribution of logarithmic link function-generated data to be very dissimilar to identity link function-generated data, whereas with a value of $c = 1$ the distributions were more similar.

The model estimation was done by Bayesian Markov chain Monte Carlo methods. Model misspecification in estimating parameters from a simulated data series with ZQ1 transformation and $c = 1$ resulted in incorrect parameter estimates, except for the ZQ2 model with the correct c parameter (fixed). A Bayesian GSARIMA model was tested on simulated data and estimated the parameters correctly.

Bayesian GSARIMA and GARIMA models were applied to malaria case count time series data of Gampaha District in Sri Lanka. Both a GSARIMA and a GARIMA model with a deterministic seasonality component were selected based on different criteria. The GARIMA model showed a lower DIC, but the GSARIMA model had a lower mean absolute relative error and needed less parameters. The Bayesian modelling allowed analysis of the posterior distributions of the fitted observations. For both models, on average, these distributions did not perfectly mirror the distribution of the residuals (although the GARIMA model did better). This is possibly an indication that the assumption of an under-laying negative binomial distribution was not entirely appropriate. Both models could account for autocorrelation in the data, although the GSARIMA model was slightly better based on the Ljung-Box statistic.

Instead of Bayesian methods, estimation using an iteratively reweighed least squares algorithm as used by Benjamin and colleagues [139] should also be feasible for GSARIMA modelling. An unpolished function for modelling GARMA models is available in the free R module VGAM [166], but this does not (yet) handle GSARIMA models.

Acknowledgements

Chapter 9 – General Discussion

This thesis brings together many aspects of malaria epidemiology in Sri Lanka, which converged towards developing forecasting models for better strategic planning. Assessment of the malaria situation and associated risk factors through incidence maps set the stage for predicting the impact of the tsunami on malaria, and the exploration of forecasting models, in a country where malaria incidence is unstable and fluctuates in intensity both spatially and temporally.

Rationale and goal of this thesis

The overall aim of this thesis was to describe the spatial and seasonal distribution of malaria in Sri Lanka and associated factors, and to develop a malaria forecasting system. Despite a wealth of high resolution data collected over decades, a malaria forecasting system was not in place, and there was a need for detailed island-wide maps of malaria incidence to permit the assessment of the malaria situation and its determinants. The Anti Malaria Campaign of Sri Lanka also needed malaria risk maps and malaria case number predictions to assist in the planning for malaria control. This need was emphasized with the adoption by the AMC of the WHO "Global Malaria Control Strategy" [22]. Early detection and containment or prevention of epidemics and the regular assessment of the malaria situation and its determinants are two of the four main elements of this strategy.

Contributions of this thesis in context

Malaria risk assessment

An online publication that resulted from the work described in Chapter 2 of this thesis provided health professionals and the larger general public in 2003 with the first island-wide incidence maps of *P. vivax* and *P. falciparum* malaria at sub district resolution in Sri Lanka. This publication also contained maps depicting seasonality and recent trends of malaria, and discussed associated risk factors. This work has been accessed over 30 000 times since publication [167], making it today the most viewed publication of the Malaria Journal since its inception. Possibly a large proportion of the accesses were by prospective visitors to Sri Lanka wishing to inform themselves about the risk of malaria and the need for prophylaxis. Maps of malaria disease distribution over the island had not been available to a wide public, until a (book) publication [2] of a map based on 1989-1994 incidence data at district resolution. The last annual report of the AMC that contained island wide malaria

maps was published in 1992 [168]. Travel medicine Internet sites described in their advice to travellers to Sri Lanka merely that the risk of malaria is present all year round in all areas (below 800 m altitude), except in the districts of Colombo, Kalutara, and Nuwara Eliya, and sometimes unrealistic maps [169] are (still) posted, showing most of the country at risk of malaria.

The tsunami that hit Sri Lanka on 26 December 2004, and the ensuing international concern about possibilities of an increase of vector borne diseases triggered the online publication of the work described in Chapter 3 of this thesis. The experiences gained during the process of writing and publishing the work of Chapter 2 allowed a rapid assessment of the risk of a malaria outbreak as a result of the tsunami, which was considered low. Thanks to the cooperation of peer reviewers and editors of the Malaria Journal, the assessment was published just one month after the disaster, making the publication timely and relevant for the planning and guidance of national and international malaria control efforts in response to the tsunami event. Also this work has reached a large audience, being the second most viewed publication of the Malaria Journal since its inception, with over 15000 access events today [167]. One year after the tsunami, the impact of the tsunami on the malaria situation and the malaria control efforts was evaluated, as is described in Chapter 4. This evaluation established that the tsunami had not affected the downward trend in malaria cases in Sri Lanka, confirming the prediction that was made immediately after the tsunami event. This was in contrast to the negative effect of the tsunami on the malaria situation in the Andaman and Nicobar Islands [89].

Development of a malaria forecasting system

A malaria forecasting system could assist in the efficient allocation of resources for malaria control, especially when malaria is unstable and fluctuates in intensity both spatially and temporally, as is the situation in Sri Lanka. Variability in malaria case numbers is influenced by factors intrinsic to malaria such as infectivity, immunity and susceptibility of vectors and humans, and extrinsic, environmental factors such as rainfall. Malaria prediction can be done both by biological process models and statistical models [129]. This thesis was restricted to statistical models. When autocorrelation is present in the data, malaria cases can be predicted based on previous / neighbouring values, and data of extrinsic covariates could improve prediction [105,131,170,171]. In Sri Lanka, relatively long and good quality time

series of monthly malaria case data are available at district resolution. A number of different types of classic time series prediction models [90], initially without extrinsic covariates, because of the extra effort involved in routinely processing of extrinsic covariates, were tested for prediction accuracy at prediction horizons of one to four months (Chapter 7). The models that gave the highest prediction accuracy were generally the class of multiplicative seasonal autoregressive integrated moving average (SARIMA) models, where seasonality was modelled through random effects. It was then explored if extrinsic covariates could improve prediction accuracy of these models.

Early malariologists had observed that particularly the centre-south-west of the island (then called Ceylon) was prone to malaria epidemics when the rainfall during the south-west monsoon period was less than usual, whereas in the north and east malaria epidemics seemed positively associated with rainfall [8]. The observation of Bouma and van der Kaay [9] that epidemics in the period from 1867 to 1943 were significantly more likely to occur during El Niño years (when the southwest monsoon is was likely to be less intense and involved less rainfall) provided hope that ENSO could be used to predict malaria epidemics. However, since the introduction of large scale vector control in 1945, the relationship between epidemics and rainfall (and ENSO) could have been weakened. Furthermore, in more recent years, the direction of the association between precipitation and ENSO has been reversed [10]. Therefore, in order to establish how rainfall could assist in malaria prediction, the present analyses focused on statistical rainfall – malaria relationships. The historic notion that malaria endemicity is negatively correlated to the spatial distribution of rainfall in Sri Lanka [7] was confirmed with statistical methods accounting for spatial autocorrelation (Chapter 2 addendum). Also, the notion that the modality of seasonality of rainfall and that of malaria are correlated in space [8] was statistically confirmed by a regression using a new seasonality index (Chapter 5), and relationships were found to be unaltered in the control era.

If malaria time series are significantly correlated to rainfall case time series, while lagging a few months behind rainfall, rainfall data could have potential use for short term malaria prediction. Some studies observed links between the two [7,8,11,108], yet others did not find a strong [109] or an obvious correlation [110,172]. In this thesis, the temporal correlation between monthly malaria case time series and monthly

rainfall time series was explored for each district separately (Chapter 6). For most districts, strong positive correlations were found for malaria time series lagging zero to three months behind rainfall. However, these results were spurious, as was shown in an analysis using prewhitening. This was the first application of the prewhitening technique, which removes autocorrelation and seasonality from the series, in malaria-rainfall cross correlation analysis. Only weak cross correlations were found in prewhitened series, thus indicating that rainfall might have little potential use in a malaria forecasting system. However, the cross correlation analyses (with and without prewhitening) had the drawbacks that (i) data had to be detrended before analysis, thus masking inter-annual effects, and (ii) the analyses did not allow for potentially seasonally varying effects of rainfall on malaria.

Subsequent inter-annual analysis [107] showed strong negative correlations between malaria and rainfall for a group of districts in the centre-west of the country, confirming the notion of early malariologists that this area was prone to drought associated epidemics [8]. By modifying the technique of inter-annual analysis to allow for seasonally varying effects, it was shown that the effect of rainfall on malaria varied according to the season (and geography). Thus, seasonally varying effects of rainfall on malaria case counts may explain weak overall cross-correlations found in prewhitened series.

With the knowledge of statistical rainfall – malaria relationships gained, it was then tested if the inclusion of rainfall as a covariate in the SARIMA models could improve prediction accuracy. The rainfall covariate was studied in linear, non linear and seasonally variable forms (Chapter 7). Only for a few districts, a modest improvement in prediction error was made when rainfall was included as a covariate. This modest improvement was not deemed to be sufficient to merit investing in a forecasting system for which rainfall data are routinely processed. Therefore, a malaria forecasting system based on SARIMA models (without extrinsic covariates) was introduced to the AMC (addendum to Chapter 7), which is being tested at the time of writing. To our knowledge, this is the first operational malaria forecasting system to include monthly forecasts based on statistical modeling.

Throughout the statistical modelling in Chapter 7, it was assumed that logarithmically transformed malaria case data were approximately Gaussian distributed. However, such an approximation is less close when case counts are close to zero, and the

prediction interval may be incorrect. Given that recent malaria case counts were low in districts in Sri Lanka, there was a need for predictive models that did not suffer from such limitations.

Although there is a growing literature on time series models for count data, no count data models with SARIMA structure were yet available. Therefore, in this thesis (Chapter 8), a class of generalised multiplicative seasonal autoregressive integrated moving average (GSARIMA) models for the parsimonious and observation-driven modelling of non Gaussian, non stationary and / or seasonal time series data was developed, by extending a class of generalized autoregressive moving average (GARMA) models [139]. The parameter estimation was done by Bayesian Markov chain Monte Carlo methods. This was demonstrated by the application of a GSARIMA model with negative binomial distribution and a logarithmic link function to a monthly time series of counts of new malaria episodes in Gampaha district in Sri Lanka. Although the Bayesian estimation requires substantial implementation skill and computing power, the developed GSARIMA models may find wide application in temporal modelling of rare diseases.

Limitations of this thesis and directions for future research

The main limitation of the work described in this thesis is the data quality, which is difficult to control for in studies that depend on external data sources. Although the data were of relatively good quality compared to those available in many tropical countries, the following data issues leading to spatial and temporal bias were recognized:

- Due to the armed conflict in the North and East, there was a shortage of trained microscopists in these areas and only a small part of the clinical cases was microscopically confirmed [25];

- The microscopically confirmed cases may be about 70% in the rest of the country [26], thus underestimating the true number of cases;

- The amount of cross checking of blood slides at central level since 1989 was limited;

- Although estimated as low, the amount of self treatment is not sufficiently known, thus it is difficult to adjust case data for this;

- Patients who seek treatment at non-governmental health facilities are not registered;

- Cases detected in occasional mass blood surveys in selected villages in high risk areas are also included in the statistics;

- Malaria infections may not have been contracted at the place of reporting;

- The population census data may be less reliable in the North and East due to armed conflict;

- Aggregated case records from the health facilities were not corrected for recrudescence of *P. falciparum* or relapse of *P. vivax*;

- Results from approximately 6000 malaria rapid diagnostic kits supplied by UN agencies [82] for use in remote areas where no microscopists are available were not included in the surveillance statistics in 2005.

Further, the number of extrinsic variables tested for potential to explain the spatial distribution of malaria was limited to rainfall, population density, altitude (as a proxy for temperature), safety clearance (as proxy for conflict), shoreline of rivers and reservoirs, and a few classes of land use cover. The extrinsic variables tested for potential to explain the temporal distribution of malaria were limited to rainfall. The performance of rainfall with distributed lags [105,127], or at thresholds, remains to be explored. Other related variables such as soil moisture content and river flow might give better results than rainfall itself, as these are more directly related to malaria vector breeding conditions.

The temporal effect of temperature remains to be studied, although it may only affect malaria in the hill country, as in the rest of the country, the temperature is always in a range conducive to malaria transmission.

An obvious extrinsic variable missing from these analyses is the amount and quality of vector control, which has proven itself to have strong impact on malaria in Sri Lanka in the past. Whereas vector control data (number of houses sprayed with insecticide) was collected by the AMC, monthly data was (only) available since around 1992 (depending on the district), and the length of the time series was considered too short for this analysis. Furthermore, vector control data was complicated by issues such as change of active ingredient, and vector resistance

development, but above all, by feedback (vector control will have an impact on malaria incidence, but malaria incidence will influence the amount of vector control). Nevertheless, it might be worthwhile to include the impact of vector control in future research, especially because it may explain long term trends in malaria incidence. With such a variable, possibly time series models would no longer need to employ a stochastic trend to render the series stationary. Another variable related to vector control that is likely to explain long term trends is conflict. In a preliminary exploratory analysis (not shown), the annual number of malaria cases was highly correlated to the number of civilians killed, missing or injured in the conflict zone. However, the quality of this data was disputed and not regarded fit for publication.

In this thesis, extrinsic variables which act on mosquito vectors were employed to explain malaria case distribution, but the distribution and abundance of vector mosquitoes themselves was not included in explanatory models. Such data might clarify the paradoxical relationships found between rainfall and malaria. Unfortunately, time series of mosquito collections were scarce in Sri Lanka and not of sufficient quality to be included in the models. Similarly, immunity was not taken into account because of the lack of time series of anti-malarial anti-body prevalence. It would be interesting to compare the results of the models employed in this thesis with models that incorporate entomological and immunological data.

Apart from the data quality, another limitation of the thesis is that the models employed do not take into account space-time interaction. Instead, spatial models were fitted to study the spatial distribution of malaria between areas / districts, and temporal models were used to study the temporal distribution of malaria within districts. A space-time malaria forecasting model for the whole of Sri Lanka could have more statistical power than a separate model for each district. Such a model should allow for regionally varying functions of covariates and take into account spatial auto correlation between districts. This limitation could be addressed in future research.

This thesis focused on models for short term monthly malaria count predictions of one to four months, and rainfall did not contribute much to model prediction. It is well possible that rainfall, or ENSO, could be useful in prediction of inter-annual malaria anomalies [118,128].

References

1. Fonseka J, Mendis KN: **A metropolitan hospital in a non-endemic area provides a sampling pool for epidemiological studies on vivax malaria in Sri Lanka.** *Trans R Soc Trop Med Hyg* 1987, **81:**360-364.

2. Konradsen F, Amerasinghe FP, Van der Hoek W, Amerasinghe PH: *Malaria in Sri Lanka: Current knowledge on transmission and control.* Colombo: International Water Management Institute; 2000.

3. Amerasinghe PH, Amerasinghe FP, Konradsen F, Fonseka KT, Wirtz RA: **Malaria vectors in a traditional dry zone village in Sri Lanka.** *Am J Trop Med Hyg* 1999, **60:**421-429.

4. Smith CE: **The significance of mosquito longevity and blood-feeding behaviour in the dynamics of arbovirus infections.** *Med Biol* 1975, **53:**288-294.

5. Craig MH, Snow RW, Le Sueur D: **A climate based distribution model of malaria transmission in sub-Saharan Africa.** *Parasitol Today* 1999, **15:**105-111.

6. Martens P, Kovats RS, Nijhof S, de Vries P, Livermore MTJ, Bradley DJ, Cox J, McMichael AJ: **Climate change and future populations at risk of malaria.** *Global Environmental Change* 1999, **9:**S89-S107.

7. Rustomjee KJ: *Observations upon the epidemiology of malaria in Ceylon.* Sessional paper XXIV-1944. Colombo: Ceylon government press; 1944.

8. Gill CA: **Some points in the epidemiology of malaria arising out of the study of the malaria epidemic in Ceylon in 1934-35.** *Trans R Soc Trop Med Hyg* 1936, **29:**427-466.

9. Bouma MJ, Van der Kaay HJ: **The El Nino Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics?** *Trop Med Int Health* 1996, **1:**86-96.

10. Zubair L, Chandimala J: **Epochal changes in ENSO-streamflow relationships in Sri Lanka.** *Journal of Hydrometeorology* 2006, **7:**1237-1246.

11. Mendis C, Gamage-Mendis AC, De Zoysa AP, Abhayawardena TA, Carter R, Herath PR, Mendis KN: **Characteristics of malaria transmission in Kataragama, Sri Lanka: a focus for immuno-epidemiological studies.** *Am J Trop Med Hyg* 1990, **42:**298-308.

12. Graves PM, Fernando D, Attanayake N: *Intensified malaria control programme in Sri Lanka with emphasis on primary care approach.* Consultants' report Washington D.C.: IDA/World Bank; 1995.

13. Mendis C, Del Giudice G, Gamage-Mendis AC, Tougne C, Pessi A, Weerasinghe S, Carter R, Mendis KN: **Anti-circumsporozoite protein antibodies measure age related exposure to malaria in Kataragama, Sri Lanka.** *Parasite Immunol* 1992, **14:**75-86.

14. Ramasamy R, Nagendran K, Ramasamy MS: **Antibodies to epitopes on merozoite and sporozoite surface antigens as serologic markers of malaria transmission: studies at a site in the dry zone of Sri Lanka.** *Am J Trop Med Hyg* 1994, **50:**537-547.

15. Smith T, Charlwood JD, Kihonda J, Mwankusye S, Billingsley P, Meuwissen J, Lyimo E, Takken W, Teuscher T, Tanner M: **Absence of seasonal variation in malaria parasitaemia in an area of intense seasonal transmission.** *Acta Trop* 1993, **54:**55-72.

16. Pascual M, Cazelles B, Bouma MJ, Chaves LF, Koelle K: **Shifting patterns: malaria dynamics and rainfall variability in an African highland.** *Proc Biol Sci* 2008, **275:**123-132.

17. Nicholls L: **Malaria and the lost cities of Ceylon.** *Indian Med Gaz* 1921, **56:**121-130.

18. Uragoda CG: *A history of medicine in Sri Lanka.* Colombo: Sri Lanka Medical Association; 1987.

19. Rajendram S, Jayewickreme SH: **Malaria in Ceylon. Part 1. The control and prevention of epidemic malaria by the residual spraying of houses with D.D.T.** *Indian J Malariol* 1951, **5:**1-73.

20. Rawlings P, Herath PR, Kelly S: **Anopheles culicifacies (Diptera: Culicidae): DDT resistance in Sri Lanka prior to and after cessation of DDT spraying.** *J Med Entomol* 1985, **22:**361-365.

21. Pinikahana J, Dixon RA: **Trends in malaria morbidity and mortality in Sri Lanka.** *Indian J Malariol* 1993, **30:**51-55.

22. **Implementation of the global malaria control strategy. Report of a WHO Study Group on the Implementation of the Global Plan of Action for Malaria Control 1993-2000.** *World Health Organ Tech Rep Ser* 1993, **839:**1-57.**:1-57.

23. Konradsen F, Van der HW, Amerasinghe FP, Mutero C, Boelee E: **Engineering and malaria control: learning from the past 100 years.** *Acta Trop* 2004, **89:**99-108.

24. Helinski ME, Hassan MM, El-Motasim WM, Malcolm CA, Knols BG, El-Sayed B: **Towards a sterile insect technique field release of Anopheles arabiensis mosquitoes in Sudan: irradiation, transportation, and field cage experimentation.** *Malar J* 2008, **7:65.:**65.

25. Reilley B, Abeyasinghe R, Pakianathar MV: **Barriers to prompt and effective treatment of malaria in northern Sri Lanka.** *Trop Med Int Health* 2002, **7:**744-749.

26. Abeysekera T, Wickremasinghe AR, Gunawardena DM, Mendis KN: **Optimizing the malaria data recording system through a study of case detection and treatment in Sri Lanka.** *Trop Med Int Health* 1997, **2:**1057-1067.

27. Konradsen F, Van der Hoek W, Amerasinghe PH, Amerasinghe FP, Fonseka KT: **Household responses to malaria and their costs: a study from rural Sri Lanka.** *Trans R Soc Trop Med Hyg* 1997, **91:**127-130.

28. Ramasamy R, Subanesan N, Wijesundere A, Fernando NK, Ramasamy MS: **Observations on malaria patients seeking treatment in hospitals in a rural and an urban area of Sri Lanka.** *Indian J Malariol* 1992, **29:**29-34.

29. *Annual administrative report of the Anti-Malaria Campaign.*Colombo: Ministry of Health; 1989.

30. Galappaththy GNL: **A study of chloroquine resistant *Plasmodium falciparum* malaria in Sri Lanka.** *MD Thesis.* Postgraduate Institute of Medicine, University of Colombo; 2003.

31. Jayawardene R: **Illness perception: social cost and coping-strategies of malaria cases.** *Soc Sci Med* 1993, **37:**1169-1176.

32. Pinikahana J: **Illness behavior and preventive behavior of the people and malaria transmission in Sri Lanka.** *Mosquito Borne Diseases Bulletin* 1993, **10:**12-20.

33. Fernando SD, Wickremasinghe AR: **The clinical and epidemiological features of childhood malaria in a moderately endemic area of Sri Lanka.** *Southeast Asian J Trop Med Public Health* 2002, **33:**671-677.

34. Handunnetti SM, Gunewardena DM, Pathirana PP, Ekanayake K, Weerasinghe S, Mendis KN: **Features of recrudescent chloroquine-resistant *Plasmodium falciparum* infections confer a survival advantage on parasites and have implications for disease control.** *Trans R Soc Trop Med Hyg* 1996, **90:**563-567.

35. Greenwood B: **The molecular epidemiology of malaria.** *Trop Med Int Health* 2002, **7:**1012-1021.

36. Fonseka J, Mendis KN: **A metropolitan hospital in a non-endemic area provides a sampling pool for epidemiological studies on vivax malaria in Sri Lanka.** *Trans R Soc Trop Med Hyg* 1987, **81:**360-364.

37. Van der Hoek W, Konradsen F, Amerasinghe PH, Perera D, Piyaratne MK, Amerasinghe FP: **Towards a risk map of malaria for Sri Lanka: the importance of house location relative to vector breeding sites.** *Int J Epidemiol* 2003, **32:**280-285.

38. Gunawardena DM, Wickremasinghe AR, Muthuwatta L, Weerasingha S, Rajakaruna J, Senanayaka T, Kotta PK, Attanayake N, Carter R, Mendis KN: **Malaria risk factors in an endemic region of Sri Lanka, and the impact and cost implications of risk factor-based interventions.** *Am J Trop Med Hyg* 1998, **58:**533-542.

39. *Annual Statistical Report 2001.* Colombo: Ceylon Tourist Board; 2002.

40. Gunawardena DM, Briët OJT, Van der Hoek W, Amerasinghe FP, Mutuwatte LP: **Mapping malaria risk in Sri Lanka.** *Proceedings of the 58th Annual Session, Sri Lanka Association of the Advancement of Science* 2002.

41. **Department of Census and Statistics - Sri Lanka** [http://www.statistics.gov.lk/census2001/database/index.asp]

42. **North East Provincial Council** [http://www.nepc.lk/]

43. **CIP Map Server** [http://gis.cip.cgiar.org/gis/data/MapServer.htm]

44. [http://www.iwmi.cgiar.org/WAtlas/atlas.htm]

45. Shahriah P. **Personal communication.** 2004.

46. **IWMI data storage pathway** [http://www.iwmidsp.org/]

47. Best N: **Bayesian ecological modelling.** In *Disease mapping and risk assessment for public health.* Edited by Edited by Lawson AB, Biggeri A, Bohning D, Lesaffre B, Viel J-F, Bertollini R. Chichester: John Wiley & Sons Ltd.; 1999:193-201.

48. Gelfand AE, Vounatsou P: **Proper multivariate conditional autoregressive models for spatial data analysis.** *Biostatistics* 2003, **4:**11-25.

49. Bernardinelli L, Clayton D, Pascutto C, Montomoli C, Ghislandi M, Songini M: **Bayesian analysis of space-time variation in disease risk.** *Stat Med* 1995, **14:**2433-2443.

50. Kleinschmidt I, Sharp B, Mueller I, Vounatsou P: **Rise in malaria incidence rates in South Africa: a small-area spatial analysis of variation in time trends.** *Am J Epidemiol* 2002, **155:**257-264.

51. Kazembe LN, Kleinschmidt I, Sharp BL: **Patterns of malaria-related hospital admissions and mortality among Malawian children: an example of spatial modelling of hospital register data.** *Malar J* 2006, **5:93.:**93.

52. Mabaso ML, Craig M, Vounatsou P, Smith T: **Towards empirical description of malaria seasonality in southern Africa: the example of Zimbabwe.** *Trop Med Int Health* 2005, **10:**909-918.

53. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A: **Bayesian measures of model complexity and fit (with discussion).** *Journal of the Royal Statistical Society B* 2002, **64:**583-640.

54. Brooks SP, Gelman A: **Alternative methods for monitoring convergence of iterative simulations.** *Journal of Computational and Graphical Statistics* 1998, **7:**434-455.

55. Lunn DJ, Thomas A, Best N, Spiegelhalter D: **WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility.** *Statistics and Computing* 2000, **10:**325-327.

56. **R.** [www.r-project.org]

57. Sachs J, Malaney P: **The economic and social burden of malaria.** *Nature* 2002, **415:**680-685.

58. **WHO Situation Report 4 January 2005** [http://w3.whosea.org/EN/Section23/Section1108/Section1835/Section1862_8 323.htm]

59. **WHO Situation Report 14 January 2005** [http://w3.whosea.org/EN/Section23/Section1108/Section1835/Section1862_8 475.htm]

60. **Tsunami Maps via FTP from IWMI** [http://www.iwmidsp.org/iwmi/info/tsunami.asp]

61. Briët OJT, Gunawardena DM, Van der Hoek W, Amerasinghe FP: **Sri Lanka malaria maps.** *Malar J* 2003, **2:**22.

62. **Refugees and Internally Displaced Repatriation and Returns to and within Sri Lanka, Statistical Summary as of 31 October 2004** [http://www.unhcr.lk/info.asp]

63. **Center for National Operations** [http://www.health.gov.lk/links/surveillance.shtml]

64. **WHO Situation Report 7 January 2005** [http://w3.whosea.org/EN/Section23/Section1108/Section1835/Section1862_8 361.htm]

65. **Reliefweb** [http://www.reliefweb.int/w/RWB.NSF/0/76c55da2c1f3a000c1256f7f004ba42 c?OpenDocument]

66. **SLMC concerned over non registration of foreign medical teams.** Daily Mirror , 3. 1-12-2005.

67. Ratnapala R, Subramaniam K, Yapabandara MG, Fernando WP: **Chloroquine resistant *Plasmodium falciparum* in Sri Lanka.** *Ceylon Med J* 1984, **29:**135-145.

68. Hapuarachchi HA, Dayanath MY, Abeysundara S, Bandara KB, Abeyewickreme W, de Silva NR: **Chloroquine resistant *falciparum* malaria**

among security forces personnel in the Northern Province of Sri Lanka. *Ceylon Med J* 2004, **49:**47-51.

69. **Editorial.** VIRAKESARI . 1-13-2005.

70. Sahu SS: **Comparative susceptibility of *Anopheles subpictus* from fresh and brackish water areas to *Plasmodium falciparum* infection.** *Acta Trop* 1998, **70:**1-7.

71. Abhayawardana TA, Wijesuriya SR, Dilrukshi RK: ***Anopheles subpictus* complex: distribution of sibling species in Sri Lanka.** *Indian J Malariol* 1996, **33:**53-60.

72. Pathinayake SW: **Investigation in to the malaria epidemic of Pasgoda DDHS area.** *Journal of the Ruhuna Clinical Society* 1997, **4:**14-18.

73. **Technical note: Malaria risk and malaria control in Asian countries affected by the tsunami disaster** [http://mosquito.who.int/docs/Asia_tsunami_malaria_risk-v1-5Jan.pdf]

74. Surendra SN, Abhayawardana TA, De Silva BG, Ramasamy R, Ramasamy MS: ***Anopheles culicifacies* Y-chromosome dimorphism indicates sibling species (B and E) with different malaria vector potential in Sri Lanka.** *Med Vet Entomol* 2000, **14:**437-440.

75. Surendra SN, De Silva BG, Srikrishnaraj KA, Ramasamy MS, Ramasamy R: **Establishment of *An. culiciafcies* sibling species E and not B as the major malaria vector in the country.** *SLAAS* 2003,180.

76. Clark TM, Flis BJ, Remold SK: **Differences in the effects of salinity on larval growth and developmental programs of a freshwater and a euryhaline mosquito species (Insecta: Diptera, Culicidae).** *J Exp Biol* 2004, **207:**2289-2295.

77. Herath PR, Joshi GP: **Pesticide selection pressure on *Anopheles subpictus* in Sri Lanka: comparison with two other Sri Lankan anophelines.** *Trans R Soc Trop Med Hyg* 1989, **83:**565-567.

78. Karunaratne SH: **Insecticide cross-resistance spectra and underlying resistance mechanisms of Sri Lankan anopheline vectors of malaria.** *Southeast Asian J Trop Med Public Health* 1999, **30:**460-469.

79. Karunaratne SH, Hemingway J: **Malathion resistance and prevalence of the malathion carboxylesterase mechanism in populations of mosquito vectors of disease in Sri Lanka.** *Bull World Health Organ* 2001, **79:**1060-1064.

80. Briët OJT, Galappaththy GNL, Konradsen F, Amerasinghe PH, Amerasinghe FP: **Maps of the Sri Lanka malaria situation preceding the tsunami and key aspects to be considered in the emergency phase and beyond.** *Malar J* 2005, **4:**8.

81. **Housing and Non Housing Building Units Damaged by the Tsunami 2004 Trincomalee District**
[http://www.statistics.gov.lk/Tsunami/final/Trincomalee/]

82. **Sri Lanka: Facts regarding post-tsunami recovery six months on**
[http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6DY5LD]

83. **Sri Lanka housing falling far short of goals**
[http://www.twincities.com/mld/pioneerpress/news/local/13787999.htm]

84. **Dengue: Stopping a potentially deadly threat in Sri Lanka**
[http://www.reliefweb.int/rw/RWB.NSF/db900SID/EKOI-6LK3UF]

85. Kelly-Hope LA, Yapabandara AM, Wickramasinghe MB, Perera MD, Karunaratne SH, Fernando WP, Abeyasinghe RR, Siyambalagoda RR, Herath PR, Galappaththy GN, Hemingway J: **Spatiotemporal distribution of insecticide resistance in Anopheles culicifacies and Anopheles subpictus in Sri Lanka.** *Trans R Soc Trop Med Hyg* 2005, **99:**751-761.

86. Balaraman K, Sabesan S, Jambulingam P, Gunasekaran K, Boopathi Doss PS: **Risk of outbreak of vector-borne diseases in the tsunami hit areas of southern India.** *Lancet Infect Dis* 2005, **5:**128-129.

87. Gunasekaran K, Jambulingam P, Srinivasan R, Sadanandane C, Doss PB, Sabesan S, Balaraman K, Das P: **Malaria receptivity in the tsunami-hit coastal villages of southern India.** *Lancet Infect Dis* 2005, **5:**531-532.

88. Villholth KG, Amerasinghe PH, Jeyakumar P, Panabokke CR, Woolley O, Weerasinghe MD, Amalraj N, Prathepaan S, Bürgi N, Sarath Lionelrathne DMD, Gamage Indrajith N, Pathirana SRK: *Tsunami Impacts on Shallow Groundwater and Associated Water Supply on the East Coast of Sri Lanka.* IWMI; 2005.[http://www.iwmi.cgiar.org/TSUNAMI/pdf/Tsunami_Impacts_on_Shallow_Groundwater.pdf]

89. Krishnamoorthy K, Jambulingam P, Natarajan R, Shriram AN, Das PK, Sehgal S: **Altered environment and risk of malaria outbreak in South Andaman, Andaman & Nicobar Islands, India affected by tsunami disaster.** *Malar J* 2005, **4:**32.

90. Chatfield C: *The analysis of time series: an introduction.* Boca Raton: Chapman & Hall / CRC; 2004.

91. Bai J, Perron P: **Estimating and Testing Linear Models with Multiple Structural Changes.** *Econometrica* 1998, **65:**47-78.

92. Mabaso ML, Craig M, Ross A, Smith T: **Environmental predictors of the seasonality of malaria transmission in Africa: the challenge.** *Am J Trop Med Hyg* 2007, **76:**33-38.

93. Gemperli A, Sogoba N, Fondjo E, Mabaso M, Bagayoko M, Briet OJ, Anderegg D, Liebe J, Smith T, Vounatsou P: **Mapping malaria transmission in West and Central Africa.** *Trop Med Int Health* 2006, **11:**1032-1046.

94. Grover-Kopec EK, Blumenthal MB, Ceccato P, Dinku T, Omumbo JA, Connor SJ: **Web-based climate information resources for malaria control in Africa.** *Malar J* 2006, **5:**38.

95. Childs DZ, Cattadori IM, Suwonkerd W, Prajakwong S, Boots M: **Spatiotemporal patterns of malaria incidence in northern Thailand.** *Trans R Soc Trop Med Hyg* 2006, **100:**623-631.

96. Briët OJT, Vounatsou P, Gunawardena DM, Galappaththy GNL, Amerasinghe PH: **Temporal correlation between malaria and rainfall in Sri Lanka.** *Malar J* 2008, **7**.

97. *Arjuna's Atlas of Sri Lanka.* Colombo: Arjuna's Consulting Co.; 1997.

98. Burman JP, Otto MC: *Outliers in time series.* Bureau of the census statistical reseach division report series SRD Research Report Number: CENSUS/SRD/RR-88114. Washington D.C.: 1988.

99. Ribeiro Jr. PJ, Diggle PJ: **geoR: A package for geostatistical analysis.** *R-News* 2001, **1**.

100. Box GEP, Jenkins GM: **Some recent advances in forecasting and control.** *Appl Statist* 1968, **17:**91-109.

101. Mwaniki KJ, Odhiambo OR, Charles W: **Seasonal time series data imputation: comparison between feedforward neural networks and parametric approaches.** *East Afr J Stat* 2005, **1:**68-83.

102. Royston P: **An extension of Shapiro and Wilk's *W* test for normality to large samples.** *Applied Statistics* 1982, **31:**115-124.

103. Moran PAP: **Notes on continuous stochastic phenomena.** *Biometrika* 1950, **37:**17-23.

104. Sun D, Tsutakawa RK, Kim H, He Z: **Spatio-temporal interaction with disease mapping.** *Stat Med* 2000, **19:**2015-2035.

105. Teklehaimanot HD, Lipsitch M, Teklehaimanot A, Schwartz J: **Weather-based prediction of *Plasmodium falciparum* malaria in epidemic-prone regions of Ethiopia I. Patterns of lagged weather effects reflect biological mechanisms.** *Malar J* 2004, **3:**41.

106. Worrall E, Connor SJ, Thomson MC: **A model to simulate the impact of timing, coverage and transmission intensity on the effectiveness of indoor residual spraying (IRS) for malaria control.** *Trop Med Int Health* 2007, **12:**75-88.

107. Craig MH, Kleinschmidt I, Nawn JB, Le Sueur D, Sharp BL: **Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: part I. The impact of climatic factors.** *Trop Med Int Health* 2004, **9:**1247-1257.

108. Clemesha WW: **Brief account of the natural history of malaria in Ceylon.** *Ceylon J Sci* 1934, **3:**157-172.

109. Van der Hoek W, Konradsen F, Perera D, Amerasinghe PH, Amerasinghe FP: **Correlation between rainfall and malaria in the dry zone of Sri Lanka.** *Ann Trop Med Parasitol* 1997, **91:**945-949.

110. De Alwis R, Wijesundere A, Ramasamy MS, Ramasamy R: **Epidemiology of malaria in Aralaganvila in the Polonnaruwa district.** In *Current status of malaria research in Sri Lanka*. Edited by Edited by Ramasamy R. Kandy: Institute of fundamental studies; 1990:80-84.

111. Briët OJT, Vounatsou P, Gunawardena DM, Galappaththy GNL, Amerasinghe PH: **Models for short term malaria prediction in Sri Lanka.** *Malar J* 2008, **7.**

112. Thomson MC, Connor SJ: *A framework for field research in Africa: Malaria early warning systems: concepts, indicators and partners.*Geneva: Roll Back Malaria Cabinet Project; 2001.

113. Briët OJT, Vounatsou P, Amerasinghe PH: **Malaria seasonality and rainfall seasonality in Sri Lanka are correlated in space.** *Geospatial Health* 2008, **2:**183-190.

114. Bi P, Tong S, Donald K, Parton KA, Ni J: **Climatic variables and transmission of malaria: a 12-year data analysis in Shuchen County, China.** *Public Health Rep* 2003, **118:**65-71.

115. Dev V, Dash AP: **Rainfall and malaria transmission in north-eastern India.** *Ann Trop Med Parasitol* 2007, **101:**457-459.

116. Singh N, Sharma VP: **Patterns of rainfall and malaria in Madhya Pradesh, central India.** *Ann Trop Med Parasitol* 2002, **96:**349-359.

117. Craig MH, Kleinschmidt I, Le Sueur D, Sharp BL: **Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: part II. The impact of non-climatic factors.** *Trop Med Int Health* 2004, **9:**1258-1266.

118. Thomson MC, Mason SJ, Phindela T, Connor SJ: **Use of rainfall and sea surface temperature monitoring for malaria early warning in Botswana.** *Am J Trop Med Hyg* 2005, **73:**214-221.

119. Wijesundera MS: **Malaria outbreaks in new foci in Sri Lanka.** *Parasitol Today* 1988, **4:**147-150.

120. Briët OJT, Vounatsou P, Amerasinghe PH: **Generalized (S)ARIMA models for count data with application to malaria time series.** Unpublished.

121. Briet OJ, Galappaththy GN, Amerasinghe PH, Konradsen F: **Malaria in Sri Lanka: one year post-tsunami.** *Malar J* 2006, **5:**42.

122. Rajakaruna RS, Amerasinghe PH, Galappaththy GNL, Konradsen F, Briët OJT, Alifrangis M: **Current situation of malaria and anti-malarial drug resistance in Sri Lanka.** *Ceylon J Sci* 2008, **37:**15-22.

123. Githeko AK, Lindsay SW, Confalonieri UE, Patz JA: **Climate change and vector-borne diseases: a regional analysis.** *Bull World Health Organ* 2000, **78:**1136-1147.

124. Basher R, Cane MA: **Climate Variability, Climate Change and Malaria.** In *The contextual determinants of malaria*. Edited by Edited by Casman E, Dowlatabadi H. Washing, DC.: Resources for the Future Press; 2002:189-215.

125. Krishnamurti TN, Chakraborti A, Mehta VM, Mehta AV: **Experimental prediction of climate-related malaria incidence.** 2007,[http://www.publichealth.pitt.edu/supercourse/SupercoursePPT/29011-30001/29021.ppt#442]

126. Meskel TG. **Personal communication.** 2007.

127. Teklehaimanot HD, Schwartz J, Teklehaimanot A, Lipsitch M: **Weather-based prediction of *Plasmodium falciparum* malaria in epidemic-prone regions of Ethiopia II. Weather-based prediction systems perform comparably to early detection systems in identifying times for interventions.** *Malar J* 2004, **19:**44.

128. Thomson MC, Doblas-Reyes FJ, Mason SJ, Hagedorn R, Connor SJ, Phindela T, Morse AP, Palmer TN: **Malaria early warnings based on seasonal climate forecasts from multi-model ensembles.** *Nature* 2006, **439:**576-579.

129. Myers MF, Rogers DJ, Cox J, Flahault A, Hay SI: **Forecasting disease risk for increased epidemic preparedness in public health.** *Adv Parasitol* 2000, **47:**309-330.

130. Abeku TA, De Vlas SJ, Borsboom G, Teklehaimanot A, Kebede A, Olana D, van Oortmarssen GJ, Habbema JD: **Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: a simple seasonal adjustment method performs best.** *Trop Med Int Health* 2002, **7:**851-857.

131. Abeku TA, De Vlas SJ, Borsboom GJ, Tadege A, Gebreyesus Y, Gebreyohannes H, Alamirew D, Seifu A, Nagelkerke NJ, Habbema JD: **Effects of meteorological factors on epidemic malaria in Ethiopia: a statistical modelling approach based on theoretical reasoning.** *Parasitology* 2004, **128:**585-593.

132. Mabaso ML, Smith T: **Empirical modelling and mapping of seasonality of malaria transmission by *Anopheles gambiae sensu lato* in Africa.** Unpublished.

133. Zubair L. **Personal communication.** 2007.

134. Mahanama S, Koster R, Zubair L, Riehl R: **On the predictability of seasonal streamflow with a Land Surface Model: Water and energy budget studies for Sri Lanka.** Unpublished.

135. Briet OJ, Galappaththy GN, Amerasinghe PH, Konradsen F: **Malaria in Sri Lanka: one year post-tsunami.** *Malar J* 2006, **5:**42.

136. King G: **Statistical Models for Political Science Event Counts: Bias in Conventional Procedures and Evidence for the Exponential Poisson Regression Model.** *American Journal of Political Science* 1988, **32:**838-863.

137. Ravines RR, Schmidt AM, Migon HS: **Revisiting distributed lag models through a Bayesian perspective.** *Applied Stochastic Models in Business & Industry* 2006, **22:**193-210.

138. Nobre AA, Schmidt AM, Lopes HF: **Spatio-temporal models for mapping the incidence of malaria in Pará.** *Environmetrics* 2005, **16:**291-304.

139. Benjamin MA, Rigby RA, Stasinopoulos DM: **Generalized Autoregressive Moving Average Models.** *Journal of the American Statistical Association* 2003, **98:**214-223.

140. West M, Harrison J: *Bayesian forecasting and dynamic models.* Springer-Verlag; 1997.

141. Gamerman D: **Markov chain Monte Carlo for dynamic generalized linear models.** *Biometrika* 1998, **85:**215-227.

142. Cox DR: **Statistical analysis of time series: some recent developments.** *Scandinavian Journal of Statistics* 1981, **8:**93-115.

143. Jackman S: *Time series models for discrete data: solutions to a problem with quantitative studies of international conflict.* Stanford: Department of Political Science, Stanford University; 1998.

144. Czado C, Kolbe A: **Model-based quantification of the volatility of options at transaction level with extended count regression models.** *Applied Stochastic Models in Business and Industry* 2007, **23:**1-21.

145. Davis RA, Dunsmuir WTM, Streett SB: **Observation-driven models for Poisson counts.** *Biometrika* 2003, **90:**777-790.

146. Jung RC, Kukuk M, Liesenfeld R: **Time series of count data: modeling, estimation and diagnostics.** *Computational Statistics & Data Analysis* 2006, **51:**2350-2364.

147. Zeger SL, Qaqish B: **Markov regression models for time series: a quasi-likelihood approach.** *Biometrics* 1988, **44:**1019-1031.

148. Grunwald G, Hyndman R, Tedesco L, Tweedie R: **Non-Gaussian conditional linear AR(1) models.** *Australian & New Zealand Journal of Statistics* 2000, **42:**479-495.

149. Brandt PT, Williams JT: **A linear Poisson autoregressive model: The PAR(p).** *Political Analysis* 2001, **9:**164-184.

150. Heinen A: *Modelling time series count data: An autoregressive conditional Poisson model.* Louvain-la-Neuve: Catholic University of Louvain; 2003.

151. McKenzie E: **Some ARMA models for dependent sequences of Poisson count.** *Advances in Applied Probability* 1988, **20:**822-835.

152. Alzaid AA, Al-Osh MA: **Some autoregressive moving average processes with generalized Poisson marginal distributions.** *Annals of the Institute of Statistical Mathematics* 1993, **45:**223-232.

153. Jung RC, Tremayne AR: **Binomial thinning models for integer time series.** *Statistical Modelling* 2006, **6:**81-96.

154. Zeger SL: **A regression model for time series of counts.** *Biometrika* 1988, **75:**621-629.

155. Mabaso ML, Vounatsou P, Midzi S, Da Silva J, Smith T: **Spatio-temporal analysis of the role of climate in inter-annual variation of malaria incidence in Zimbabwe.** *Int J Health Geogr* 2006, **5:**20.

156. Knorr-Held L, Besag J: **Modelling risk from a disease in time and space.** *Stat Med* 1998, **17:**2045-2060.

157. Waller LA, Carlin BP, Xia H, Gelfand AE: **Hierarchical spatio-temporal mapping of disease rates.** *Journal of the American Statistical Association* 1997, **92:**607-617.

158. Box GEP, Jenkins GM: *Time series analysis, forecasting and control.* San Fransisco: Holden-Day; 1970.

159. Jones MC: **Randomly choosing parameters from the stationary and invertibility region of autoregressive-moving average models.** *Applied Statistics* 1987, **36:**134-138.

160. Congdon P: **Time series models for discrete variables.** In *Bayesian Models for Categorical Data*. John Wiley & Sons Canada, Ltd; 2001:289-319.

161. Sturtz S, Ligges U, Gelman A: **R2WinBUGS: A Package for Running WinBUGS from R.** *Journal of Statistical Software* 2005, **12:**1-16.

162. **gsarima: Two functions for Generalized SARIMA time series simulation** [http://cran.r-project.org/]

163. Dunn PK, Smyth GK: **Randomized quantile residuals.** *Journal of Computational and Graphical Statistics* 1996, **5:**236-244.

164. Ljung GM, Box GEP: **On a measure of lack of fit in time series models.** *Biometrika* 1978, **65:**297-303.

165. Chen Y-T: **On the robustness of Ljung-Box and McLeod-Li Q tests: a simulation study.** *Economics Bulletin* 2002, **3:**1-10.

166. Yee TW, Wild CJ: **Vector generalized additive models.** *Journal of the Royal Statistical Society, Series B, Methodological* 1996, **58:**481-493.

167. **Malaria Journal list of most viewed articles since its inception** [http://www.malariajournal.com/mostviewedalltime]

168. *Annual administrative report of the Anti-Malaria Campaign.*Colombo: Ministry of Health; 1992.

169. **Detailkarte Malaria Risikogebiete Sri Lanka** [http://www.fit-for-travel.de/reisemedizin/reiseziele/index.jsp]

170. Kleinschmidt I, Vounatsou P, Sharp B, Curtis B, Hay S, Cox J: **Space-time modelling of small area malaria incidence in relation to remote-sensed inter-annual climatic variation.** In *Spatial statistical analysis, modelling and mapping of malaria in Africa, PhD Thesis I. Kleinschmidt*. 2001:107-122.

171. Abeku TA, De Vlas SJ, Borsboom G, Teklehaimanot A, Kebede A, Olana D, van Oortmarssen GJ, Habbema JD: **Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: a simple seasonal adjustment method performs best.** *Trop Med Int Health* 2002, **7:**851-857.

172. Ramasamy MS, Kulasekera R, Srikrishnaraj KA, Ramasamy R: **Population dynamics of anthropophilic mosquitoes during the northeast monsoon season in the malaria epidemic zone of Sri Lanka.** *Med Vet Entomol* 1994, **8:**265-274.

Appendix
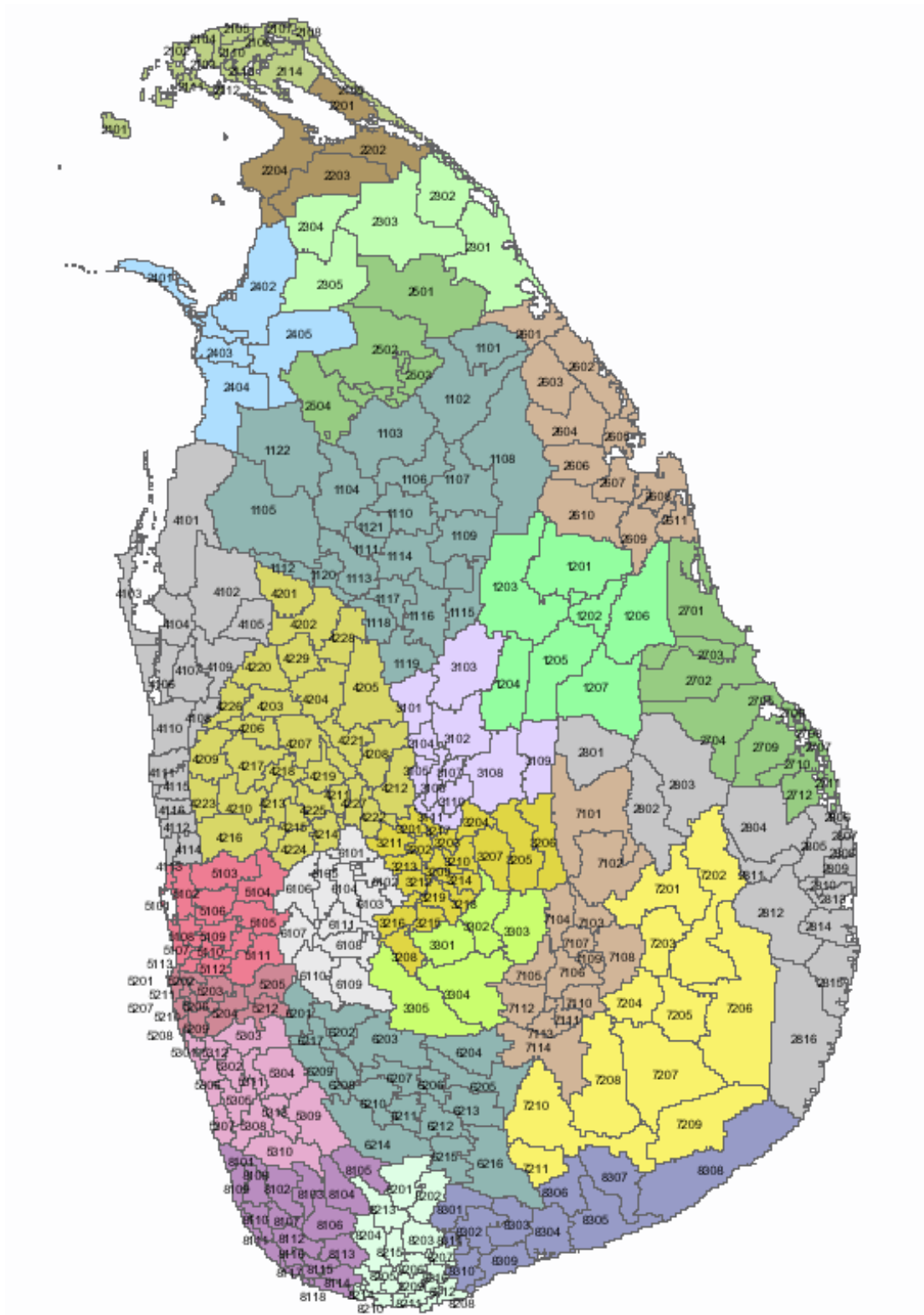
Chapter 2 Additional file 1a – Divisional Secretariat Divisions

Map of Divisional Secretariat Divisions in Sri Lanka. Source: The Survey Department, Sri Lanka.
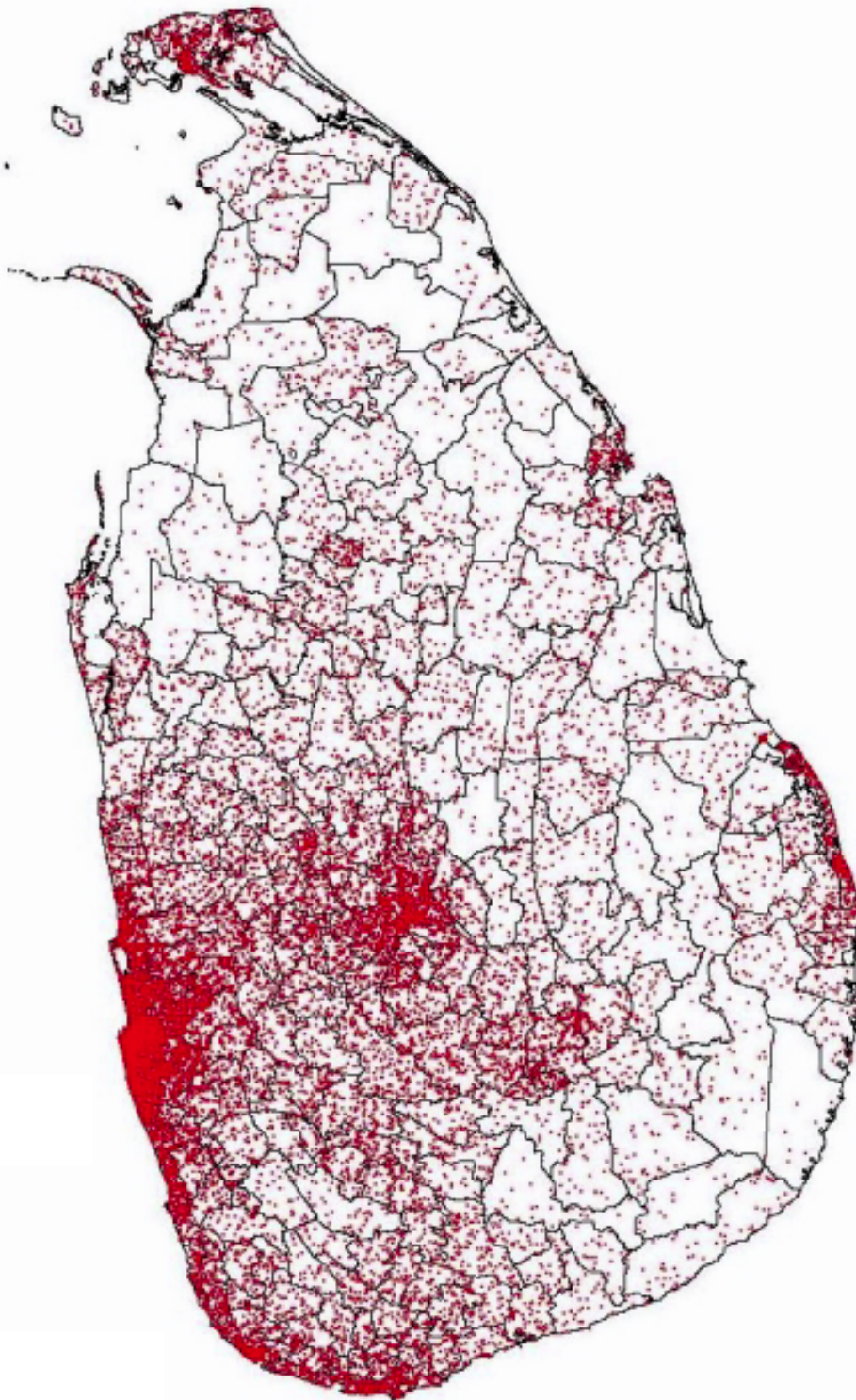
Chapter 2 Additional File 1b

Divisional Secretariat Divisions. List of Divisional Secretariat Division names and their number corresponding to the number printed in additional file 1a.

ID	DSD	DISTRICT	ID	DSD	DISTRICT
1101	Padawiya	Anuradhapura	2304	Tunukkai	Mullaitivu
1102	Kebitigollewa	Anuradhapura	2305	Pandiyankulam (mantai east)	Mullaitivu
1103	Medawachchiya	Anuradhapura	2401	Mannar	Mannar
1104	Nuwaragam palatha central	Anuradhapura	2402	Mantai west	Mannar
1105	Nochchiyagama	Anuradhapura	2403	Nanaddan	Mannar
1106	Rambewa	Anuradhapura	2404	Musali	Mannar
1107	Kahatagasdigiliya	Anuradhapura	2405	Madhu	Mannar
1108	Horowupothana	Anuradhapura	2501	Vavuniya north	Vavuniya
1109	Galenbindunuwewa	Anuradhapura	2502	Vavuniya	Vavuniya
1110	Mihintale	Anuradhapura	2503	Vavuniya south	Vavuniya
1111	Nachchaduwa	Anuradhapura	2504	Vengala cheddikulam	Vavuniya
1112	Rajanganaya	Anuradhapura	2601	Padawi sripura	Trincomalee
1113	Talawa	Anuradhapura	2602	Kuchchaveli	Trincomalee
1114	Tirappane	Anuradhapura	2603	Gomarankadawala	Trincomalee
1115	Palugaswewa	Anuradhapura	2604	Morawewa	Trincomalee
1116	Kekirawa	Anuradhapura	2605	Trincomalee town & gravets	Trincomalee
1117	Ipalogama	Anuradhapura	2606	Thampalagamuwa	Trincomalee
1118	Galnewa	Anuradhapura	2607	Kinniya	Trincomalee
1119	Palagala	Anuradhapura	2608	Muttur	Trincomalee
1120	Thambuththegama	Anuradhapura	2609	Seruwila	Trincomalee
1121	Nuwaragam palatha east	Anuradhapura	2610	Kantale	Trincomalee
1122	Mahawilachchiya	Anuradhapura	2611	Verugal eachchalampattu	Trincomalee
1201	Medirigiriya	Polonnaruwa	2701	Koralai pattu north	Batticaloa
1202	Lankapura	Polonnaruwa	2702	Koralai pattu	Batticaloa
1203	Hingurakgoda	Polonnaruwa	2703	Koralai pattu west	Batticaloa
1204	Elahera	Polonnaruwa	2704	Eravur pattu	Batticaloa
1205	Thamankaduwa	Polonnaruwa	2705	Eravurpattu town	Batticaloa
1206	Welikanda	Polonnaruwa	2706	Manmunai north	Batticaloa
1207	Dimbulagala	Polonnaruwa	2707	Manmunai pattu	Batticaloa
2101	Delft	Jaffna	2708	Kattankudi	Batticaloa
2102	Kaytes	Jaffna	2709	manmunai west	Batticaloa
2103	Chankanai	Jaffna	2710	Manmunai south west	Batticaloa
2104	Sandilippai	Jaffna	2711	Manmunai south & eruvil	Batticaloa
2105	Tellippalai	Jaffna	2712	Porativu pattu	Batticaloa
2106	Kopai	Jaffna	2801	Dehiattakandiya	Ampara
2107	Karaveddy	Jaffna	2802	Padiyatalawa	Ampara
2108	Pointpedro	Jaffna	2803	Maha-oya	Ampara
2109	Maruthankerny	Jaffna	2804	Uhana	Ampara
2110	Uduvil ' chunnakam '	Jaffna	2805	Sammanturai	Ampara
2111	Velanai	Jaffna	2806	Kalmunai	Ampara
2112	Jaffna	Jaffna	2807	Karativu	Ampara
2113	Nallur	Jaffna	2808	Nainativu	Ampara
2114	Chavakachcheri	Jaffna	2809	Attalachena	Ampara
2201	Pachchilaipallai	Kilinochchi	2810	Akkaraipattu	Ampara
2202	Kandavalai	Kilinochchi	2811	Ampara	Ampara
2203	Karachchi	Kilinochchi	2812	Damana	Ampara
2204	Pooneryn	Kilinochchi	2813	Alayadivembu	Ampara
2301	Maritime pattu	Mullaitivu	2814	Thirukkovil	Ampara
2302	Puthukkudiyiruppu	Mullaitivu	2815	Potuvil	Ampara
2303	Oddusuddan	Mullaitivu	2816	Lahugala	Ampara

ID	DSD	DISTRICT	ID	DSD	DISTRICT
3101	Galewela	Matale	4204	Mahawa	Kurunegala
3102	Naula	Matale	4205	Polpithigama	Kurunegala
3103	Dambulla	Matale	4206	Kobeigane	Kurunegala
3104	Pallepola	Matale	4207	Wariyapola	Kurunegala
3105	Yatawatta	Matale	4208	Ibbagamuwa	Kurunegala
3106	Matale	Matale	4209	Bingiriya	Kurunegala
3107	Ambanganga	Matale	4210	Kuliyapitiya west	Kurunegala
3108	Laggala pallegama	Matale	4211	Kurunegala	Kurunegala
3109	Wilgamuwa	Matale	4212	Ridigama	Kurunegala
3110	Raththota	Matale	4213	Kuliyapitiya east	Kurunegala
3111	Ukuwela	Matale	4214	Polgahawela	Kurunegala
3201	Pujapitiya	Kandy	4215	Narammala	Kurunegala
3202	Harispattuwa	Kandy	4216	Pannala	Kurunegala
3203	Patha dumbara	Kandy	4217	Panduwasnuwara	Kurunegala
3204	Panwila	Kandy	4218	Katupotha	Kurunegala
3205	Uda dumbara	Kandy	4219	Maspotha	Kurunegala
3206	Minipe	Kandy	4220	Kotawehera	Kurunegala
3207	Meda dumbara	Kandy	4221	Ganewattaa	Kurunegala
3208	Pasbage korale	Kandy	4222	Mawathagama	Kurunegala
3209	Mahanuwara four gravets	Kandy	4223	Udubaddawa	Kurunegala
3210	Kundasale	Kandy	4224	Alawwa	Kurunegala
3211	Tumpane	Kandy	4225	Weerambugedara	Kurunegala
3212	Udunuwara	Kandy	4226	Rasnayakapura	Kurunegala
3213	Yatinuwara	Kandy	4227	Mallawapitiya	Kurunegala
3214	Patha hewaheta	Kandy	4228	Ehetuwewa	Kurunegala
3215	Uda palatha	Kandy	4229	Ambanpola	Kurunegala
3216	Ganga ihala korale	Kandy	5101	Negambo	Gampaha
3217	Akurana	Kandy	5102	Katana	Gampaha
3218	Delthota	Kandy	5103	Divulapitiya	Gampaha
3219	Doluwa	Kandy	5104	Meerigama	Gampaha
3301	Kotmale	Nuwaraeliya	5105	Attanagalla	Gampaha
3302	Hanguranketha	Nuwaraeliya	5106	Minuwangoda	Gampaha
3303	Walapane	Nuwaraeliya	5107	Wattala	Gampaha
3304	Nuwaraeliya	Nuwaraeliya	5108	Ja-ela	Gampaha
3305	Ambangamuwa korale	Nuwaraeliya	5109	Gampaha	Gampaha
4101	Vanathavillu	Puttalam	5110	Mahara	Gampaha
4102	Karuwalasgawewa	Puttalam	5111	Dompe	Gampaha
4103	Kalpitiya	Puttalam	5112	Biyagama	Gampaha
4104	Puttalam	Puttalam	5113	Kelaniya	Gampaha
4105	Nawagaththegama	Puttalam	5201	Colombo	Colombo
4106	Mundala	Puttalam	5202	Kolonnawa	Colombo
4107	Mahakumbukkadawala	Puttalam	5203	Kaduwela	Colombo
4108	Pallama	Puttalam	5204	Homagama	Colombo
4109	Anamaduwa	Puttalam	5205	Hanwella	Colombo
4110	Arachchikattuwa	Puttalam	5206	Maharagama	Colombo
4111	Chilaw	Puttalam	5207	Nugegoda	Colombo
4112	Nattandiya	Puttalam	5208	Moratuwa	Colombo
4113	Wennappuwa	Puttalam	5209	Kesbewa	Colombo
4114	Dankotuwa	Puttalam	5210	Dehiwala- mount laviniya	Colombo
4115	Madampe	Puttalam	5211	Thimbirigasyaya	Colombo
4116	Mahawewa	Puttalam	5212	Padukka	Colombo
4201	Giribawa	Kurunegala	5301	Panadura	Kalutara
4202	Galgamuwa	Kurunegala	5302	Bandaragama	Kalutara
4203	Nikaweratiya	Kurunegala	5303	Horana	Kalutara

ID	DSD	DISTRICT	ID	DSD	DISTRICT
5304	Bulathsinhala	Kalutara	7203	Medagama	Monaragala
5305	Dodangoda	Kalutara	7204	Badalkumbura	Monaragala
5306	Kalutara	Kalutara	7205	Monaragala	Monaragala
5307	Beruwala	Kalutara	7206	Siyambalanduwa	Monaragala
5308	Mathugama	Kalutara	7207	Buttala	Monaragala
5309	Agalawatta	Kalutara	7208	Wellawaya	Monaragala
5310	Walallawita	Kalutara	7209	Katharagama	Monaragala
5311	Madurawela	Kalutara	7210	Thanamalwila	Monaragala
5312	Millaniya	Kalutara	7211	Sevanagala	Monaragala
5313	Palinda nuwara	Kalutara	8101	Benthota	Galle
6101	Rambukkana	Kegalle	8102	Elpitiya	Galle
6102	Mawanella	Kegalle	8103	Niyagama	Galle
6103	Aranayake	Kegalle	8104	Tawalama	Galle
6104	Kegalle	Kegalle	8105	Neluwa	Galle
6105	Galigamuwa	Kegalle	8106	Nagoda	Galle
6106	Warakapola	Kegalle	8107	Welivitiya-divithura	Galle
6107	Ruwanwella	Kegalle	8108	Karandeniya	Galle
6108	Yatiyantota	Kegalle	8109	Balapitiya	Galle
6109	Deraniyagala	Kegalle	8110	Ambalangoda	Galle
6110	Dehiowita	Kegalle	8111	Hikkaduwa	Galle
6111	Bulathkohupitiya	Kegalle	8112	Baddegama	Galle
6201	Eheliyagoda	Ratnapura	8113	Yakkalamulla	Galle
6202	Kuruwita	Ratnapura	8114	Imaduwa	Galle
6203	Ratnapura	Ratnapura	8115	Akmeemana	Galle
6204	Imbulpe	Ratnapura	8116	Bope-poddala	Galle
6205	Balangoda	Ratnapura	8117	Galle	Galle
6206	Opanayake	Ratnapura	8118	Habaraduwa	Galle
6207	Pelmadulla	Ratnapura	8201	Kotapola	Matara
6208	Elapatha	Ratnapura	8202	Pasgoda	Matara
6209	Ayagama	Ratnapura	8203	Mulatiyana	Matara
6210	Nivithigala	Ratnapura	8204	Akuressa	Matara
6211	Kahawatta	Ratnapura	8205	Malimbada	Matara
6212	Godakawela	Ratnapura	8206	Kamburupitiya	Matara
6213	Weligepola	Ratnapura	8207	Hakmana	Matara
6214	Kalawana	Ratnapura	8208	Dikwella	Matara
6215	Kolonne korale	Ratnapura	8209	Thihagoda	Matara
6216	Embilipitiya	Ratnapura	8210	Weligama	Matara
6217	Kiriella	Ratnapura	8211	Matara	Matara
7101	Mahiyangana	Badulla	8212	Devinuwara	Matara
7102	Ridimaliyadda	Badulla	8213	Pitabeddara	Matara
7103	Meegahakivula	Badulla	8214	Welipitiya	Matara
7104	Kandeketiya	Badulla	8215	Athuraliya	Matara
7105	Uva paranagama	Badulla	8216	Kirinda-puhulwell	Matara
7106	Hali-ela	Badulla	8301	Katuwana	Hambanthota
7107	Soranathota	Badulla	8302	Weeraketiya	Hambanthota
7108	Passara	Badulla	8303	Anguinukolapelessa	Hambanthota
7109	Badulla	Badulla	8304	Ambalathota	Hambanthota
7110	Ella	Badulla	8305	Hambanthota	Hambanthota
7111	Bandarawela	Badulla	8306	Suriyawewa	Hambanthota
7112	Welimada	Badulla	8307	Lunugamwehera	Hambanthota
7113	Haputale	Badulla	8308	Tissamaharama	Hambanthota
7114	Haldummulla	Badulla	8309	Tangalle	Hambanthota
7201	Bibile	Monaragala	8310	Beliatta	Hambanthota
7202	Madulla	Monaragala	8311	Okewela	Hambanthota

Chapter 2 Additional file 2 – Population

Map of population by divisional secretariat division in Sri Lanka. One dot represents 1000 people. Sources: Department of Census and Statistics (http://www.statistics.gov.lk/) and North East Provincial Council (http://www.nepc.lk/index.htm).

Chapter 2 Additional file 3 – Temperature

Maps of monthly average temperature in Sri Lanka. Source: IWMI World Water and Climate Atlas (http://www.iwmi.cgiar.org/WAtlas/atlas.htm).

Chapter 2 Additional file 4 – Altitude

Map of altitude in Sri Lanka. Colour shading as recommended by the GLOBE project (except white background). Green contour line: 500m, black contour lines: 1000m. Source: GLOBE project (http://www.ngdc.noaa.gov/seg/topo/globe.shtml).

Chapter 2 Additional file 5 – Rainfall

Maps of monthly rainfall (75% probability) in Sri Lanka. Source: IWMI World

Water and Climate Atlas (http://www.iwmi.cgiar.org/WAtlas/atlas.htm).

Chapter 2 Additional file 6 – Rivers and lakes

Map of river network and lakes / tanks in Sri Lanka. Source: The Survey Department,
Sri Lanka.

Chapter 4 additional file 1

Short/Medium Term Plan for prevention and control of possible malaria outbreaks in tsunami affected areas of Sri Lanka.

Prepared by the Anti Malaria Campaign Directorate on 29th December 2004.

Important: Please bear in mind that nearly all the tsunami affected districts of the country excluding the coastal districts from Matara to Gampaha belong to malarious areas. Therefore in the absence of clinical features suggestive of acute respiratory infections or acute gastro enteritis please consider the possibility of malaria in any febrile patient or in any patient with a history of fever.

Measures to be taken by Regional Malaria Officers of the Anti Malaria Campaign in tsunami affected districts and adjacent districts where camps for the displaced are located:

1.   Ensure availability of a buffer stock of at least 50000 tablets each of chloroquine and primaquine.

2.   Make available malaria diagnostic facilities (microscopy or Rapid Diagnostic Test Kits) to hospitals, camps for the displaced and to locations where outbreaks of malaria are suspected.

3.   Take immediate measures to introduce long lasting insecticide treated mosquito nets to all occupants of camps for the displaced in high risk malaria districts of Jaffna, Mullaitivu, Kilinochchi, Trincomalee, Batticaloa, Kalmunai (Ampara coastal), Hambantota. Collect all LLINs available with GFATM partner organizations and distribute to the camps as soon as possible. Seek partner organization support for this activity if possible.

4.   Carry out indoor residual spraying of <u>all camps for the displaced in all affected districts</u> wit insecticide on a priority basis.

5.   Ensure that measures are taken to locate new camps for displaced persons away from malaria risk areas of the district or ensure that no malaria vector breeding sites are in close proximity to camps for displaced persons. Take measures to modify/manipulate existing breeding sites to make them unfavourable to malaria vector breeding. Consider the application of larvicides or introduction of larvivorous fish.

6.   Make arrangements with adjacent inland districts or with Anti Malaria Campaign Headquarters to have standby staff such as Public Health Inspectors, Public Health Laboratory Technicians and Public Health field Officers for deployment in tsunami-affected districts for malaria control activities.

7.   Establish communication with all foreign and local medical teams working in the area and obtain reports of suspected number of patients treated for malaria. Inform all foreign medical teams regarding National Guidelines for malaria treatment and provide copies of documents to all teams.

8.   Send fortnightly (every two weeks) a report regarding activities carried out, number of malaria cases reported, number of suspected malaria patients treated with anti malarials. Inform Director Anti Malaria regarding any outbreak of fever or malaria outbreaks immediately by telephone (Tel no. 0112588947).

Chapter 4 additional file 2 (one_year-post-tsunami_additionalfile2_1475-2875-5-42-S1.rtf)

Non-exhaustive list of reported antimalarial support by non-governmental organizations (source: Reliefweb).

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RURI-6LYPJC?OpenDocument

Source: International Federation of Red Cross And Red Crescent Societies (IFRC)

Date: 13 Feb 2006

South Asia: Earthquake & Tsunamis Third/Fourth Quarterly Report Appeal No.28/2004 Operation Update No. 58

Distribution of mosquito nets in Ampara district (with ministry of health) together with appropriate health education

http://www.reliefweb.int/rw/rwb.nsf/db900SID/EVOD-6LUECU?OpenDocument

Source: Humanitarian Situation Report - Sri Lanka: 03 - 09 Feb 2006

United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 09 Feb 2006

IOM has distributed over 700 bags of rice, 500 hurricane lamps, 300 packets of tea, 15 tents, and 37 mosquito nets to families displaced from transitional shelter sites in Trincomalee district due to the prevalent political situation

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6LJ8PE?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 26 Jan 2006

Humanitarian Situation Report - Sri Lanka: 21 - 26 Jan 2006

"A total of 744 bed-nets and 50 baby kits in Valachenai, Batticaloa in collaboration with Regional Malaria Officer."

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6KWA69?OpenDocument

Source: MADRE (Madre)

Date: 10 Jan 2006

The tsunami: One year later

Supplies for pregnant women and new mothers, including infant mosquito nets and mats

http://www.reliefweb.int/rw/rwb.nsf/db900SID/ETOA-6KW3CX?OpenDocument

Source: The Salvation Army

Date: 30 Nov 2005

East Asian Tsunami Recovery Report - Nov 2005

Sector: Nutrition, health or medical services; Objectives: Provide medical exams/referral, establish health clinics, distribute mosquito nets; Beneficiaries to Date: 1,502.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6KH3JL?OpenDocument

Source: GOAL

Date: 23 Dec 2005

What GOAL has achieved in Sri Lanka after the tsunami one year on

distributing mosquito nets

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6KG5HB?OpenDocument

Source: Concern

Date: 26 Dec 2005

Concern Sri Lanka - One year on

Project Galle 2005 was established in response to the tsunami by a group of Sri Lankan and international volunteers living in the Galle district. The group identified families and distributed 7,376 family kits including essential relief items. Family kits included a mosquito net, a sleeping mat, water collection vessel, and essential cooking and eating utensils. Personal hygiene items were also added.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6KK5CS?OpenDocument

Source: United Nations Children's Fund (UNICEF)

Date: 22 Dec 2005

UNICEF Situation Report Sri Lanka 27 Nov - 22 Dec 2005

In Kilinochchi, assistance to people affected by the October floods in Kilinochchi and Mullaitivu has continued during the reporting period. UNICEF responded by distributing tarpaulins, 1000 liter water tanks, water pumps, sleeping mats, mosquito nets

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6KG4GG?OpenDocument

Source: International Federation of Red Cross And Red Crescent Societies (IFRC)

Date: 15 Dec 2005

Tsunami operation - Facts and figures updated 15 Dec 2005

More than 300,000 people have received distributions of relief goods including food, cooking supplies, stoves, hygiene items, mosquito nets, lamps, clothes, sleeping mats, school uniforms, schoolbags, stationery and clothes.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6K53TS?OpenDocument

Source: Direct Relief International

Date: 15 Dec 2005

Tsunami grant summaries

The Tropical and Environmental Diseases and Health Association (TEDHA) was founded in Hikkaduwa, Sri Lanka to address environmental concerns in the country. Comprised of environmental health specialists, an epidemiologist, parasitologists, and public health inspectors, the organization has worked closely with the Ministry of Health on a national malaria control program. Following the tsunami, TEDHA initiated voluntary tsunami relief health activities in Thotagamuwa-Hikkaduwa, including a vector control assessment of the area. This assessment confirmed a strong need for vector control measures in the area, especially in relief camps.

With a grant from Direct Relief, TEDHA has implemented vector borne disease prevention programs in Hikkaduwa and Hambantota, both located on the southern coast of Sri Lanka, and both devastated by the tsunami. Hikkaduwa was severely affected by the tsunami, suffering approximately half of the Galle District's nearly 5,000 deaths. Hambantota District sustained over 5,000 causalities making it one of the worst affected districts. For survivors, a lack of housing, an unsafe water supply, limited nutritional supply, and poor hygiene conditions contribute to an increased risk

of communicable diseases such as diarrhea, dengue, malaria, filariasis, and Japanese encephalitis.

In Hikkaduwa and Hambantota, TEDHA has distributed 18,000 insecticide treated mosquito nets, procured by Direct Relief, to families residing in displaced persons camps and affected neighborhoods. Recipients receive training on use of the nets and on measures to effectively prevent disease. Bed nets will be retreated as needed by TEDHA's community health volunteers who will regularly monitor net use.

In addition, Direct Relief provided 5,000 insecticide-treated mosquito nets, with a value of $30,457, to Sarvodaya to assist in their (separate) vector control programs in relief camps.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6K53H7?OpenDocument

Source: Direct Relief International

Date: 15 Dec 2005

One year after tsunami, Direct Relief International remains committed to providing vital medical resources

Over 168,800 families in Sri Lanka and India are being protected from malaria and other insect-borne diseases through the provision of 170,100 insecticide-treated mosquito nets and fogging devices;

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6K375J?OpenDocument

Source: Church World Service (CWS)

Date: 14 Dec 2005

Tsunami anniversary: In global agency's largest natural disaster response, aid workers see signs of recovery at people level

In Sri Lanka, Church World Service's Pakistan-Afghanistan regional staff provided food and water, tents, mats, sheets, mosquito nets, health supplies, kitchen utensils, clothing and medicine to some 56,100 families.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KOCA-6JZJ9G?OpenDocument

Source: Medical Emergency Relief International (Merlin)

Date: 12 Dec 2005

Preventing disease outbreaks

Thousands of people have been living in makeshift camps which are overcrowded and mosquito-ridden. In these difficult conditions, maintaining standards of personal hygiene and cleanliness are paramount if people are to stay healthy. Merlin has been helping to prevent disease outbreaks in camps by distributing more than 15,000 hygiene kits, organising clean-ups and providing camp cleaning equipment. Merlin has also trained more than 800 volunteers to promote good hygiene practices in camps.

Merlin's project is reaching more than 1 million people in total, including 3,000 mothers who have been given essential items for their newborn babies, and 120,000 people who have received mosquito nets. Two hundred midwives have also been provided with emergency delivery equipment and supplies from Merlin.

Over the coming months, Merlin will construct and equip seven permanent health centres to replace those that were destroyed. These facilities together served more than 100,000 people before the tsunami. Merlin is also helping to strengthen the existing health systems, for example, by training health workers, improving laboratories and helping to develop an emergency operational plan for disease outbreaks.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/DPAS-6JPF3V?OpenDocument

Source: Adventist Development and Relief Agency International (ADRA)

Date: 01 Dec 2005

South Asia: ADRA publishes report on tsunami response and launches remembrance campaign

Provision of mosquito nets.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6K88CA?OpenDocument

Source: Swiss Agency for Development and Cooperation (SDC)

Date: 30 Nov 2005

Two projects in Sri Lanka: Reconstruction of schools and houses

The first SDC/HA measure was the shipment of 150 truckloads of locally acquired relief goods (including mats, covers, mosquito nets, water cans, cooking sets, soap,

toiletries, candles, matches) for around 3,000 homeless families in Matara district. 250 water tanks were organized and set up to ensure a supply of drinking water in Matara.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6KS3AJ?OpenDocument

Source: Government of Canada

Date: 24 Nov 2005

12,000 mosquito nets and floor mats were distributed

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6HH76M?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 21 Oct 2005

Humanitarian Situation Report - Sri Lanka: 15 - 21 Oct 2005

Dengue alerts are ongoing in Matara district and among aid workers the OCHA Galle field office reports. Two cases of dengue have been reported by the Spanish Red Cross and Caritas International. IOM Sri Lanka is in the process of developing an information campaign aimed at delivering messages about the high tendency for an outbreak of dengue and malaria due to increased breeding sites following the seasonal rains. In Trincomalee, the Medical Health Officer with the support of IOM launched an awareness raising campaign on dengue and malaria prevention, which will be aired through the local cable TV network. This announcement will be telecast in local languages, initially for seven continuous days and then every Sunday through December 2005.

The broadcast message states: "There is potential risk for outbreaks of dengue and malaria in Trincomalee again this season. Let us protect ourselves from these diseases by destroying mosquito breeding sites such as bottles, coconut shells, polythene bags, plastic containers and other places where water can stagnate."

In Batticaloa, as part of the prevention exercise for dengue and malaria, an environmental hygiene programme has been developed by the CHSO (community health surveillance officer) to inform residents on how to destroy breeding sites for vectors of the diseases.

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 29 Sep 2005

Humanitarian Situation Report - Sri Lanka: 23 - 29 Sep 2005

In Trincomalee, an anti-malaria campaign and vector surveillance was recently conducted by health officials in Kinniya and Trincomalee town. Four malaria cases have been reported from Kinniya. UNICEF distributed 1,000 insecticide treated nets in Kinniya. Furthermore, a dengue fever awareness programme is being conducted by 100 community volunteers trained by the Ministry of Health, every Saturday in Trincomalee. Public warning messages are being issued to prevent an outbreak of dengue fever.

Source: United Nations Children's Fund (UNICEF)

Date: 28 Jun 2005

UNICEF Situation Report Sri Lanka 28 Jun 2005

66,000 families served with kits including hygiene materials, mosquito nets, lanterns, cooking utensils, buckets, water purification tablets, clothes, mattresses and sleeping bags.

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 16 Jun 2005

Sri Lanka: Facts regarding post-tsunami recovery six months on

- 101 emergency health kits have been provided to hospitals and clinics by UNICEF and WHO benefiting some 1,500,000 tsunami-affected people

- Approx. 6000 malaria rapid diagnostic kits and over 100,000 anti-malarial tablets supplied by UN agencies.

- 48,000 impregnated mosquito nets have been provided by UNICEF and WHO with 50,000 more are on the way.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/EVOD-6CYHRE?OpenDocument&cc=lka

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 16 Jun 2005

Sri Lanka: Facts regarding post-tsunami recovery six months on

Five-thousand mosquito nets were handed over to the Deputy Director of Health Services in Ampara by UNICEF to support the Anti-Malaria Campaign. Another 5,400 nets were provided to the Health Sevice in Jaffna district. An increase in mosquito-borne diseases is likely due to the seasonal North East monsoon rains.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6DV3DH?OpenDocument

Source: Lutheran World Relief (LWR)

Date: 30 Jun 2005

Six months later: a tsunami update from Lutheran World Relief

Provision of mosquito nets.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/SODA-6DS8J9?OpenDocument

Source: Medair

Date: 27 Jun 2005

Tsunami crisis: 6 month review of Medair activities in Sri Lanka

Essential relief items distributed to 2,300 families including; buckets, jerry cans, mosquito nets, mats, milk powder, soap, sachets of water purification chemicals.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/EVOD-6DCGME?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 14 Jun 2005

IOM/OCHA Sri Lanka: Humanitarian Situation Report 10 - 14 Jun 2005

The situation in Mandana camp, Thirukkovil division, Ampara district provides a good example of how the government, UN agencies and NGOs remain pro-active in their response to continuing humanitarian relief concerns. As recently as 9 June, the population of the Mandana camp, which held 592 families in March, was down to 161

families. Construction of transitional shelters, lack of transportation and incidents of violence had caused some residents to relocate, but the principal reason for the exodus was a report of several cases of Hepatitis A which quickly evolved into a rumour of the existence of a yellow fever outbreak, prompting a camp-wide scare. UN agency and NGO representatives with government health authorities took swift preventive measures, including safeguarding water supplies, spraying for mosquitoes and stepping up awareness raising activities regarding Hepatitis A and a campaign to assure residents there was no evidence of yellow fever.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/EVOD-6CYHRE?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 02 Jun 2005

Humanitarian Situation Report - Sri Lanka: 27 May - 2 June 2005

Five-thousand mosquito nets were handed over to the Deputy Director of Health Services in Ampara by UNICEF to support the Anti-Malaria Campaign. Another 5,400 nets were provided to the Health Sevice in Jaffna district. An increase in mosquito-borne diseases is likely due to the seasonal North East monsoon rains.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/EVOD-6CPHQ5?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 24 May 2005

IOM/OCHA Sri Lanka: Humanitarian Situation Report 20 - 24 May 2005

WHO is providing fogging machines for mosquito control against malaria and dengue to a number of Deputy Provincial Directors of Health (DPDH) Services in the districts. In addition, WHO has recently conducted a fogging machine operation and maintenance workshop in Matara district for new operators of the fogging equipment and for Public Health Inspectors from Galle, Matara and Hambantota districts.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/VBOL-6CJHBG?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 19 May 2005

Humanitarian Situation Report - Sri Lanka: 13 - 19 May 2005

Some 12,400 mosquito nets, which were handed over by UNICEF to the Deputy Director of Health Services in Batticaloa district, are being distributed to families in the malaria prevalent areas of the district. An increase in Mosquito-borne diseases is likely due to the seasonal North East monsoon rains.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/LSGZ-6CNBNN?OpenDocument

Source: Oxfam

Date: 04 May 2005

Tsunami crisis - situation update: Providing shelter, rebuilding livelihoods

In Trincomalee, a temporary mobile shed for public performance was made and erected in Narasima Malai camp. In Kuncahavalli a drama on diarrhoea, hygiene and mosquito borne diseases was performed to 100 people living in Narasima Mali camp and 75 people from Alnuriya Vidiyalayam camp.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/MHII-6BZ5MX?OpenDocument

Source: American Red Cross

Date: 27 Apr 2005

Red Cross spurs anti-malaria campaign

By Alice Kociejowski and Stacey M. Winston, special to Redcross.org

Wednesday, April 27, 2005 -- Colombo, Sri Lanka -- It has been four months since the tsunami, and the International Red Cross/Red Crescent Movement continues to provide invaluable assistance to vulnerable families -- notably it began an anti-malaria campaign.

At the request of the Ministry of Health, the Sri Lankan Red Cross (SLRCS) and the International Federation of Red Cross and Red Crescent Societies (Federation), in cooperation with the International Committee of Red Cross (ICRC), are distributing treated mosquito nets across the Ampara district to tsunami affected and indirectly affected families, as part of a nationwide anti-malaria campaign.

Ampara has one of the highest incidence rates of malaria in Sri Lanka, and mosquito nets, if used properly, provide simple but effective prevention against this deadly disease.

"Malaria can be a difficult disease to control, especially in a post-disaster situation, but mosquito nets are one of the easiest and most effective solutions," said Jeff Chinn, American Red Cross relief team member.

As the rainy season approaches and the risk of vector borne diseases increases, distribution of mosquito nets becomes more essential. Since the tsunami struck, the SLRCS and the Federation have distributed over 66,000 nets in Galle, Matara, Hambantota and Ampara.

This distribution is just a small, but vital part of the Red Cross work that aims to improve the lives of vulnerable people in all areas of Sri Lanka.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6JF83M?OpenDocument

Source: International Committee of the Red Cross (ICRC)

Date: 04 Apr 2005

Tsunami disaster in Sri Lanka : The response of the International Committee of the Red Cross (ICRC) 4 Apr 2005

Working with the local authorities, the Sri Lankan Red Cross and the ICRC have delivered over 35,000 family kits to welfare centres and transit camps and other displaced persons in the north and east of the country. Such kits typically contain floor mats, bed sheets, soap, towels, buckets, jerry cans and plastic dishes. In addition, over 100 welfare centres in the same regions were provided with cooking pots and utensils for communal cooking. Nearly 12,000 pieces of clothing, 22,000 blankets, 11,000 kitchen sets and 3,300 kerosene lamps have been distributed to displaced families. The ICRC is assisting the Canadian, Japanese, Swiss, Austrian and American Red Cross Societies to provide 30,000 displaced families with monthly hygiene kits over a six-month period. The kits contain soap, toothpaste, sanitary towels, bath towels, mosquito coils, etc.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/KHII-6B53VW?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 02 Apr 2005

Indonesia, Sri Lanka, Thailand: Earthquake and Tsunami OCHA Situation Report No. 34

In Ampara, the local health service is coordinating the distribution of 70,618 insecticide-treated mosquito nets supplied by several agencies. The total mosquito net requirement for the district is 83,802.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/VBOL-6AZHDG?OpenDocument

Source: United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

Date: 31 Mar 2005

Humanitarian Situation Report - Sri Lanka: 25 - 31 Mar 2005

In Ampara, the local health service is coordinating the distribution of insecticide-treated mosquito nets by a number of agencies including: UNICEF, 20,000 nets, GOAL, 20,000, Merlin, 2,450, MSF, 3,900, ICRC, 5,000, LIONS, 15,000 and Medair 4,268. The total mosquito net requirement for the district is 83,802.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6AS2SF?OpenDocument

Source: Caritas

Date: 24 Mar 2005

Caritas reviews Sri Lanka programmes - three months after tsunami

In Galle division mosquito nets have also been distributed to families in temporary shelters as malaria and dengue fever are endemic to the coastal areas of Sri Lanka.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/EGUA-6ARQF6?OpenDocument

Source: Oxfam

Date: 23 Mar 2005

OI Tsunami External Bulletin #24 of 23 Mar 2005

In Matara and Hambantota, a distribution of non-food-relief items was successfully completed, and local women and men received 10,834 hygiene items, including nappies, hot water flasks, soap, mosquito nets, washing bowls, feeding cups and spoons, a cleaning brush for bottles, sponge, sheets and a pillow. Family packs were also distributed.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/VBOL-6AQJW6?OpenDocument

Source: International Organization for Migration (IOM)

Date: 21 Mar 2005

IOM Sri Lanka: Tsunami response program update 21 Mar 2005

IOM distributed 326 mosquito nets in two IDP camps in Eachchilampattu DS division on behalf of Trincomalee Lions Club.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6AL8UV?OpenDocument

Source: Medair

Date: 17 Mar 2005

South Asia Emergency - Operations update

Distribution of essential relief items to 2,294 families including; 2,247 buckets, 2,161 jerry cans, 4,258 mosquito nets, 2,114 mats, 2,088 boxes of milk powder, 1070 bars of soap, 77,580 sachets of water purification chemicals.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/RMOI-6AL4PE?OpenDocument

Source: HelpAge International

Date: 17 Mar 2005

After the tsunami: Latest from Sri Lanka

HelpAge Sri Lanka is working with four local community organisations in Ampara, Batticaloa and Trincomalee districts to distribute non-food items such as pots, pans, mosquito nets, soap, disinfectants, spoons, knives, plates, and other items required to cope with daily needs.

http://www.reliefweb.int/rw/rwb.nsf/db900SID/VBOL-6B8ERY?OpenDocument

Source: ZOA Refugee Care

Date: 08 Feb 2005

Update on ZOA tsunami relief work in Sri Lanka - 8 Feb 2005

Dispatching of the first of 90,000 high quality mosquito nets commenced (distribution to be completed before the end of June).

Chapter 8 appendix

```
library(R2WinBUGS)
model310101x<-function(){
      ##Priors##
      #beta ~ dflat() #model with external variable
      beta <- 0 #model without external variable
      phi.star1 ~ dunif(-0.999, 0.999) #model with a first order
seasonal autoregressive parameter
#     phi.star1 <- 0 #model without a first order seasonal
autoregressive parameter
#     theta.star1 ~ dunif(-0.999, 0.999) #model with a first order
seasonal moving average parameter
      theta.star1 <- 0 #model without a first order seasonal moving
average parameter
      psi ~ dgamma(0.01,0.01)
      for (i in 1:p){
            alpha.db[i] <- round(0.5*(i+1)-0.01)
            beta.db[i] <- round(0.5*i+1-0.01)
            r[i] ~ dbeta(alpha.db[i], beta.db[i])
            r.map[i] <- 2*r[i]-1
      }
      y.phi[1,1] <- r.map[1]
      for (i in 2:p){
            for (k in 1:(i-1)){
            y.phi[k,i] <- y.phi[k,i-1] -r.map[i]*y.phi[i-k,i-1]
            }
            y.phi[i,i] <- r.map[i]
      }
      phi[1] <- 0 #use this option if phi[1] is omitted
#     phi[1] <- y.phi[1,p] #use this option if phi[1] is included
      phi[2] <- 0 #use this option if phi[2] is omitted
#     phi[2] <- y.phi[2,p] #use this option if phi[2] is included
      phi[3] <- y.phi[3,p]
      ####

      ##Error for first w observations##
      for (t in 1: w){
            u[t] <- 0
      }
      ####

      ##Likelihood ##
      for (t in (w+1):N){
            y[t] ~ dnegbin(pr[t],psi)
            pr[t] <- psi/(psi+lambda[t])
            lambda[t] <- exp(m[t])
            m[t] <- beta * x[t] +log(max(c,cut(y[t-1]))) -beta*x[t-1]
+sum(AR[ ,t]) +sum(ARSAR[ ,t]) +SAR[t] +(theta.star1)*cut(u[t-12])
            for (k in 1:p){
                  AR[k,t] <-
 phi[k]*log(max(c,cut(y[t-k])))
-phi[k]*beta*x[t-k]
-phi[k]*log(max(c,cut(y[t-k-1])))
+phi[k]*beta*x[t-k-1]
                  ARSAR[k,t]<-
-phi.star1*phi[k]*log(max(c,cut(y[t-k-12])))
+phi.star1*phi[k]*beta*x[t-k-12]
+phi.star1*phi[k]*log(max(c,cut(y[t-k-1-12])))
-phi.star1*phi[k]*beta*x[t-k-1-12]
            }
```

```
            SAR[t]<-
 phi.star1*log(max(c,cut(y[t-12])))
-phi.star1*beta*x[t-12]
-phi.star1*log(max(c,cut(y[t-1-12])))
+phi.star1 *beta*x[t-1-12]
            u[t] <- log(max(c,cut(y[t]))/cut(lambda[t]))
            LL[t] <- cut(psi)*log(cut(pr[t])) +cut(y[t])*log(1-
cut(pr[t])) +loggam(cut(y[t]) +cut(psi)) -loggam(cut(y[t])+1) -
loggam(cut(psi))
      }
      cut.psi <- cut(psi)
      ####

      ##Deviance##
      Dev <- -2*sum(LL[(w +1) : N])
      ####
}
write.model(model310101x, con = "model310101x.txt")

inits<- list(list(psi=10, phi.star1=0))
parameters<-c("psi", "phi", "phi.star1", "Dev")
data<-list(N=408, w=16, p=3, c=1, y= y[1:408], x=x[1:408])

#estimation:
winbugs.output <- bugs(data, inits, parameters,
    model.file="model310101x.txt",
    n.iter=11000, n.burnin=1000, n.thin=1, n.chains=1, debug=T,
    bugs.directory = "c:/Program Files/WinBUGS14/",
    working.directory = NULL, clearWD = TRUE)
```

List of publications (chronological)

Briët OJT (1996) Human odours as attractant for *Culex quinquefasciatus* (Diptera: Culicidae) Report 4996, Department of Entomology, Wageningen Agricultural University

Briët OJT (1997) Refractory genes of *Anopheles stephensi* Liston (Diptera: Culicidae) to *Plasmodium falciparum* (Haemosporida: Plasmodiidae). IPO DLO/ Wageningen Agricultural University, thesis

Briët OJT (1998) Comparing CDC-light traps to Human landing catches in Mali. Project Thesis part 1, Department Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine.

Briët OJT (1998) Study on mosquito population growth and parity rates. Project Thesis part 2 Department Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine.

Braks MAH, Koenraadt CJM, Briët OJT, Takken W (1999) Behavioural responses of *Culex quinquefasciatus* (Diptera: Culicidae) to human odour, carbon dioxide and lactic acid. *Proceedings of the section Experimental and Applied Entomology of the Netherlands Entomological Society (NEV)*, Amsterdam **10**, 91-98.

Sissoko MS, Briët OJT, Sissoko M *et al.* (2000) The impact of irrigated rice cultivation on the incidence of malaria in children in the region of Niono. *Cahiers Agricultures* **9**, 427.

Dolo G, Briët OJT, Dao A *et al.* (2000) The relationship between rice cultivation and malaria transmission in the irrigated Sahel of Mali, West Africa. *Cahiers Agricultures* **9**, 425.

Kleinschmidt I, Omumbo J, Briët OJT *et al.* (2001) An empirical malaria distribution map for West Africa. *Tropical Medicine and International Health,* 2001 Oct;6(10):779-786

Gunawardena DM, Briët OJT, Amerasinghe FP, Van der Hoek W, Muthuwatta L (2002) Mapping malaria risk in Sri Lanka. Proceedings of the third MIM Pan African Conference on Malaria, Arusha, November 17 – 22, 2002 Abstract 320. 195 http://www.mim.su.se/english/events/3rd_mim_conf/abstracts.html

Briët OJT (2002) A simple method for calculating mosquito mortality rates correcting for seasonal variations in recruitment. *Medical and Veterinary Entomology*, 16: 22-27

Briët OJT, Dossou-Yovo J, Akodo E, van de Giesen N, & Teuscher T (2002) The relationship between *Anopheles gambiae* density and rice cultivation in the savannah zone and forest zone of Côte d'Ivoire. Wim van der Hoek, editor (2002) Malaria and Agriculture. Abstracts for the SIMA Special Seminar at the 3rd MIM Pan-African Conference on Malaria, Arusha, Tanzania, 20 November 2002. SIMA Document 4. Colombo, Sri Lanka: International Water Management Institute. http://www.iwmi.cgiar.org/sima/FILES/word/sima%20doc%204.doc

Gunawardena DM, Briët OJT, van der Hoek W, Amerasinghe FP and Muttuwatta LP (2002). Planning malaria control interventions in an endemic area of Sri Lanka using GIS. In:Proceedings of the Spatial Information Research Centre's 14th Colloquium (Rigby J, Skelly C & Whigham PA, eds.), Wellington, New Zealand

Briët OJT, Henry MC, Dossou Yovo J, Geneau R, Akodo, E, Audibert M (2003) Single or Double Cropping of Rice in the Savannah Zone of Côte d'Ivoire: Impacts on Malaria. International forum on ecosystems approaches to human health, Montreal, May 18 – 23, 2003
http://www.idrc.ca/ev.php?URL_ID=24029&URL_DO=DO_TOPIC

Briët OJT, Dossou-Yovo J, Akodo E, van de Giesen N, & Teuscher T (2003) The relationship between *Anopheles gambiae* density and rice cultivation in the savanna zone and forest zone of Côte d'Ivoire. *Tropical Medicine and International Health* **8**, 439-448

Briët OJT, Gunawardena DM, van der Hoek W & Amerasinghe FP (2003) Sri Lanka malaria maps. Malaria Journal **2**, 22

Dolo G, Briët OJT, Dao A *et al.* (2004) Malaria transmission in relation to rice cultivation in the irrigated Sahel of Mali. *Acta Tropica* **89**, 147-159.

Sissoko MS, Dicko A, Briët OJT, Sissoko M *et al.* (2004) Malaria incidence in relation to rice cultivation in the irrigated Sahel of Mali. *Acta Tropica* **89**, 161-170.

Yapi YG, Briët OJT, Diabate S *et al.* (2005) Rice irrigation and schistosomiasis in savanna and forest areas of Côte d'Ivoire. *Acta Tropica*, **93**, 201-211.

Trung HD, Bortel WV, Sochantha T, Keokenchanh K, Briet OJ, Coosemans M (2005) Behavioural heterogeneity of Anopheles species in ecologically different localities in Southeast Asia: a challenge for vector control. *Tropical Medicine and International Health* **10**, 251-26.

Briët OJT, Galappaththy GNL, Konradsen F and Amerasinghe PH, Amerasinghe FP (2005) Maps of the Sri Lanka malaria situation preceding the tsunami and key aspects to be considered in the emergency phase and beyond. Malaria Journal **4**:8

Briët OJT, Galappaththy GNL, Konradsen F and Amerasinghe PH, Amerasinghe FP (2005) Maps of the Sri Lanka malaria situation preceding the tsunami and key aspects to be considered in the emergency phase and beyond [MIM-OB-252000]. The Fourth Mim Pan-African Malaria Conference 2005, New Strategies against an Ancient Scourge, Nov. 13 - 18, 2005, Yaoundé, Cameroon. MIM-2005 Abstract book, Abstract 480C. p332-333

Yapi YG, Briët OJT, Vounatsou P. (2006) Prevalence of geohelminths in savanna and forest areas of Côte d'Ivoire. *West African Journal of Medicine,* **25**, 124-125.

Briët OJT, Galappaththy GNL, Konradsen F and Amerasinghe PH (2006) Malaria in Sri Lanka: one year post-tsunami. Malaria Journal **5**, 42

Briët OJT, Galappaththy GNL, Zubair L, Vounatsou P (2006) Rainfall as predictor of malaria in Sri Lanka. European Congress of Epidemiology, Epidemiology and Health Care Practice, June 28-July 1 2006, Utrecht, The Netherlands. European Journal of Epidemiology, Vol 21 (Supplement) p 45 (abstract 247).

Gemperli A, Sogoba N, Fondjo E, Mabaso M, Bagayoko M, Briet OJT, Anderegg D, Liebe J, Smith T, Vounatsou P (2006) Mapping malaria transmission in West and Central Africa. *Tropical Medicine and International Health* **11:**1032-1046

Briët OJT (2006) Prediction of malaria in Sri Lanka using ARIMA models and rainfall. Water, Environment and Health Research in Sri Lanka, A Symposium and Workshop to Commemorate the Work of Dr. Felix Prashantha Amerasinghe (1948-2005), 28-29 August 2006, Colombo, Sri Lanka.

Rajakaruna RS, Amerasinghe PH, Galappaththy GNL, Konradsen F, Briët OJT, Alifrangis M (2008) Current situation of malaria and anti-malarial drug resistance in Sri Lanka. Ceylon J Sci, **37**(1), 15-22.

Briët OJT, Gunawardena DM, Galappaththy GNL, Amerasinghe PH, Vounatsou P (2008) Models for short term malaria prediction in Sri Lanka. Malaria J, **7**.

Briët OJT, Gunawardena DM, Galappaththy GNL, Amerasinghe PH, Vounatsou P (2008) Temporal correlation between malaria and rainfall in Sri Lanka. Malaria J, **7.**

Briët OJT, Vounatsou P, Amerasinghe PH: Malaria seasonality and rainfall seasonality in Sri Lanka are correlated in space. Geospatial Health, **2** (2), 183-190.

Briët OJT, Vounatsou P, Amerasinghe PH: Generalized (S)ARIMA models for count data with application to malaria time series. Unpublished.

Curriculum vitae

Personalia

Name: Briët

Given names: Olivier Johan Tavai

Date of birth: 22 October 1972

Place of birth: Apia

Country of birth: Samoa

Sex: Male

Nationality: Dutch

Email: o.briet@gmail.com

Work experience

Oct. 2007 – Apr. 2008:   IWMI Research Associate, International Water Management Institute, Colombo, Sri Lanka.

Oct. 2004 – Oct. 2007:   Regional Researcher, International Water Management Institute, Colombo, Sri Lanka.

Oct. 2001 – Oct. 2004:   Associate Expert Medical Entomology, Theme Water, Health & Environment, International Water Management Institute, Colombo, Sri Lanka.

Jul. 2001 – Oct. 2001: Consultant at the Prince Leopold Institute of Tropical Medicine for the performance of: Statistical analysis of the data obtained in the framework of research project on identification and characterisation of malaria vectors in SE Asia.

Feb. 1999 – Mar. 2001: Associate Professional Officer in Entomology/Ecology for the Netherlands Minister for Development Cooperation (DGIS) posted at the West Africa Rice Development Association (WARDA) Ecosystems and Health Research Consortium.

Oct.1998 – Jan. 1999: Consultant for the World Health Organisation for the performance of: A bibliographic review of technical literature on the association between rice production systems and vector-borne diseases over the period 1990-1998

University education

Apr. 2004 – present day: PhD candidate, Swiss Tropical Institute, University of Basel

Apr. 1998: GIS package training (Course Arc/info) at the Department of Geographical Information systems and Remote Sensing (GIRS) of Wageningen Agricultural University

Sept.1997 - Sept. 1998: MSc Biology and Control of Disease Vectors
London School of Hygiene and Tropical Medicine

Sept. 1992 - Aug. 1997: MSc and BSc Plant Breeding and Crop Protection
Wageningen Agricultural University
Specialisation Ecological Crop Protection

Sept. 1991 - Sept. 1992:   Environmental Studies

Warren Wilson College, Asheville, North Carolina, USA

MSc Thesis projects

Jun. 1998 - Sept. 1998:   Research on comparison of CDC-light traps to Human landing catches in Mali, and study on mosquito population growth and parity rates.

Feb. 1997 - Aug. 1997:   Practical period Entomology (21 weeks): AFLP-mapping of a resistance gene specific for *Plasmodium falciparum* in the malaria vector *Anopheles stephensi*.

Jan. 1997:   Nijmegen Catholic University: Epidemiology of infectious diseases (4 weeks):

Model building and simulation of onchocerciasis.

Jun. 1996 - Nov. 1996:   Thesis entomology (21 weeks): Human odours as attractant for *Culex quinquefasciatus*.