MALARIA MODELING & SURVEILLANCE

BENCHMARK REPORT

Submitted by
Richard Kiang
NASA Goddard Space Flight Center

June 30, 2009

EXECUTIVE SUMMARY

Malaria has been with the human race since ancient times. Nowadays, most of the tropical and subtropical countries are endemic with malaria. There are approximately 300–500 million cases worldwide and at least 1 million deaths in any given year. The falciparum malaria, which has become resistant to multiple drugs in most malarious areas, may become malignant and fatal without supportive care. The vivax and malariae malaria, although less virulent, may relapse and prolong morbidity. Before any effective vaccines become available, approximately 40% of the world's population is at risk.

Malaria transmission depends on the diverse factors that influence the vectors, parasites, human hosts, and the interactions among them. These factors may include, among others, meteorological and environmental condition, the innate and adapted immunity of the human hosts, the resistance of the vector species to *Plasmodium* infection, public health system, housing standards, vector control, road construction, irrigation projects, population movements, and military conflicts. The most apparent determinants are the meteorological and environmental parameters, such as rainfall, temperature, humidity, and vegetation. When other parameters are stable, the meteorological and environmental conditions can indeed be considered the driving factors. Remote sensing has been used in recent years for developing malaria early warning systems, particularly for Africa. For countries on which we have focused – Thailand, Indonesia and Afghanistan – there have been few studies to examine the dependency of malaria cases on these factors.

In an endemic area, the local adult population may acquire sufficient immunity after repeated infections. The disease could be deadly, however, to young children, pregnant women, those with depressed immunoresponse, and people new to the area. Because malaria is virtually nonexistent in the U.S., Americans traveling abroad and U.S. oversea forces are particularly vulnerable.

The decision support system which the Malaria Modeling and Surveillance (MMS) Project is aimed to enhance is the Global Situational Awareness Tool (GSAT), a multi-purpose system at the Air Force Special Operations Command (AFSOC). Initially developed as a civil engineering tool, it has gradually expanded to include other functionalities, including disease risk assessments and prevention for U.S. forces. In addition, we also have research partnership with a number of U.S. and foreign organizations, including the Armed Forces Institute of Medical Sciences in Thailand, the Naval Medical Research Unit-2 in Indonesia, and the public health organizations in these countries. In pursuing the MMS Project, we rely on the research partners for epidemiological and entomological data, malaria and vector surveillance, and field work support. Outputs from the MMS Project are shared with the decision support as well as the research partners.

The goal of the MMS project is to use NASA data, model outputs, and analytical and modeling expertise to enhance the decision support capabilities for malaria risk assessment and control. The capabilities which we have developed concern detection,

prediction and reduction of malaria risk. Our specific objectives are: 1) identification of potential larval habitats for major malaria vector species; 2) estimation of current and prediction of future malaria risks; and 3) estimation of spatio-temporal transmission characteristics for cost-effective malaria control.

The NASA data used in this project include those from MODIS, TRMM, AVHRR, SRTM and SIESIP. The geophysical parameters provided by these sensors include precipitation, land surface temperature, NDVI, dew point, elevation, etc. Relative humidity is also used for modeling. Although not a product, it can be computed from other parameters. Future climate conditions needed for modeling are based on output from NASA climate prediction models. A small amount of commercial high-resolution imagery data, such as those from Ikonos, was also used for our study sites.

We have developed a number of analytical methods and models to support the objectives of the MMS Project. For example, we use textural-contextual methods to detect small larval habitats, neural network methods for predicting malaria risks, and agent-based Monte Carlo model for simulating the spatio-temporal transmission of malaria and testing transmission hypotheses. The rationales for developing these models are briefly explained here.

The breeding sites of malaria mosquito species come in all shapes and sizes. The sites could be as large as a rice field for *Anopheles maculatus* and *sawadwongpori*, or as small as a small stream for *An. minimus*. In general, the breeding sites are small. Because of shadowing effects, it is more difficult to identify ground subjects when high spatial resolution data is used. In such cases, using spectral information alone for object classification would result in abysmal classification accuracy. Toward this end, we have developed classification techniques that utilize textural, contextual and other spatial characteristics for detecting ground objects.

For modeling malaria incidence from environmental and epidemiological time series, the main technique we have used in this project is the neural network method. Neural network is a vital part of machine or artificial intelligence, which is a discipline to study machine's ability for learning and adaptation, and exhibition of intelligent behaviors. The method is most useful when there is no clear mathematical relationship between dependant and independent variables. As the relationship between malaria transmission and the environment is complex and also depends on other parameters, neural network method is highly suitable for modeling malaria risk.

For testing transmission hypothesis, we have developed an agent-based discrete event simulation model which takes a considerable number of intrinsic and extrinsic parameters into account. The output of this model is the spatio-temporal distribution of malaria cases. As it is difficult to eliminate malaria once a region becomes endemic, this model would help identify the key factors that maintain malaria endemicity and the most effective way for reducing malaria burdens. In the report, the capabilities developed in the MMS Project on malaria risk detection, prediction and reduction are explained, and examples from the three countries – Thailand, Indonesia and Pakistan – are given.

GSAT consists of both classified and unclassified components. The unclassified component for disease risk assessment is partially based on some other systems or uses their output. Overall, GSAT's capabilities for disease risk assessment are on a par with the contemporary, state-of-the-art systems. We use a number of contemporary systems to illustrate the characteristics of the state-of-the-art capabilities for disease risk assessments. Examples are drawn from the work of WHO-UNICEF RBM, WHO EMRO, Dynamic Technology, NOAA, University of London School of Tropical Medicine and Hygiene, and Columbia University IRI's MEWS Net. A typical characteristic of such systems is that the malaria risk assessment is expressed in qualitative terms – such as high, medium, low, or free of risks. On the other hands, the malaria risks from our MMS Project are quantitative and expressed in number of malaria cases. As a comparison between a qualitative system and a quantitative system can only be qualitative, discussions in this report is necessarily qualitative.

Because these systems' risk assessments are qualitative in nature, few systems can readily be validated quantitatively. On the other hand, the MMS Project's risk assessments are expressed in malaria cases, and hence can be validated through hindcasting. Hindcasting, also called retrospective forecasting, is an established method in quantitative analysis for evaluating the performance of predictive models, and is especially popular in Earth science. In the context of malaria risk assessment, hindcasting involves using past environmental and meteorological data to predict past malaria case. By comparing the prediction with the reported cases, the performance of the models can easily be validated.

In addition, because information on disease prevalence (including that for malaria) is difficult to obtain, state-of-the-art malaria early warning systems often use the climate suitability for malaria transmission to represent malaria risk. Logically, climate suitability is a *necessary condition* for malaria transmission, but not a *sufficient condition*. Because climate suitability is logically and fundamentally different from disease prevalence, such representation could be misleading and the assessments from such systems must be treated cautiously, especially when users are unaware of the implication of climate suitability. In addition, climate suitability based reasoning may result in a static risk which does not change with time or season. On the other hand, our evidence based approach utilizes past epidemiological data and gives dynamic risk, which shows variation from month to month, and from season to season. In the report, the differences between climate suitability and actual disease distributions are elucidated with the fact that there are few autochthonous malaria cases in the U.S., even the climate is perfectly suitable for malaria transmission.

In sum, for the risk prediction capability alone, what distinguish our MMS Project from other state-of-the-art malaria early warning systems are: 1) quantitative vs. qualitative risks, 2) dynamic vs. static risks, 3) validated vs. hypothetical risks, and 4) integrated vs. climate suitability risks. Our project can therefore provide more accurate, more precise risk assessments at a verifiable confidence level. Finally, we point out that the capability for testing malaria transmission hypothesis in our MMS Project is not available in any

contemporary malaria early warning systems. Transmission hypothesis testing allows the most cost-effective approach be used for curtailing transmission.

On the other hand, GSAT incorporates a wide range of ancillary information and data for its decision making. Fusing together classified information and the general malaria risk assessment capabilities allows additional layers of analyses to be performed, and will enhance its risk assessment capabilities. Hence our conclusion based on a limited comparison with its general *unclassified* functionality may under represent its capability and should not be construed as inadequacy of GSAT.

Contents

EXECUTIVE SUMMARY **2**

1. INTRODUCTION **7**

1.1 THE MALARIA PROBLEM 7
1.2 WHY THE MILITARY NEEDS MALARIA MODELING AND EARLY WARNING 7
1.3 PURPOSE OF THE BENCHMARK REPORT 8

2. SUMMARY OF SYSTEMS ENGINEERING ACTIVITIES **9**

2.1 THE GOAL OF THE MMS PROJECT 9
2.2 REMOTE SENSING DATA OR RESULTS USED IN THE PROJECT 9
2.3 THE INTEGRATED SYSTEM SOLUTION DIAGRAM 10
2.4 SUMMARY OF THE EVALUATION EFFORTS 11
2.5 SIGNIFICANCE OF THE COUNTRIES STUDIED IN THE MMS PROJECT 12
2.6 SUMMARY OF THE IMPLEMENTATION EFFORTS 13
2.6.1 Textural-contextual classification methods *13*
2.6.2 Neural network methods *13*
2.6.3 Dynamic transmission model *14*
2.7 SUMMARY OF THE V&V EFFORTS 14
2.7.1 Risk detection *15*
2.7.2 Risk prediction *16*
2.7.3 Risk reduction *17*

3. BENCHMARKING **19**

3.1 QUANTITATIVE DYNAMIC VS. QUALITATIVE STATIC RISKS 19
3.1.1 World malaria risk in WHO 2005 report *19*
3.1.2 World malaria risk in WHO 2008 report *20*
3.1.3 Afghanistan malaria risk in WHO EMRO 2009 report *20*
3.1.4 Malaria risk map from a NOAA project *21*
3.2 INTEGRATED VS. CLIMATE SUITABILITY RISKS 22
3.2.1 The meaning of climate suitability *22*
3.2.2 Transmission suitability for malaria in Afghanistan *22*
3.2.3 Malaria risk map from Dynamic Technology *23*
3.2.4 Malaria risk map from MEWS Net *24*
3.2.5 What is missing in climate-suitability based reasoning *24*
3.2.6 Why the US does not have malaria *26*
3.3 VALIDATED VS. HYPOTHETICAL RISKS 27

4. CONCLUSIONS **27**

ACRONYMS **29**

REFERENCES **30**

1. INTRODUCTION

1.1 The malaria problem

Malaria has been with the human race since ancient times. The name of the disease was derived from *mala aria*, or *bad air* in Italian. In history, both in the West and the East, it was long thought that the air around swamps was bad for the health and could cause illness. People learned by experience that the bad air problem would disappear by draining the swamps. It was not until the late 18th Century that it was discovered that it was actually the mosquitoes around swamps, and not the air, that caused diseases.

The world missed the opportunity for eradicating malaria before DDT was banned. Nowadays, most of the tropical and subtropical countries are endemic with malaria. There are approximately 300–500 million cases worldwide and at least 1 million deaths in any given year. The falciparum malaria, which has become resistant to multiple drugs in most malarious areas, may become malignant and fatal without supportive care. The vivax and malariae malaria, although less virulent, may relapse and prolong morbidity. Before any effective vaccines become available, approximately 40% of the world's population is at risk.

Malaria transmission depends on the diverse factors that influence the vectors, parasites, human hosts, and the interactions among them. These factors may include, among others, meteorological and environmental condition, the innate and adapted immunity of the human hosts, the resistance of the vector species to *Plasmodium* infection, public health system, housing standards, vector control, road construction, irrigation projects, population movements, and war-like conditions. The most apparent determinants are the meteorological and environmental parameters, such as rainfall, temperature, humidity, and vegetation. When other parameters remain more or less constant, the meteorological and environmental conditions can indeed be considered the driving factors. Remote sensing has been used in recent years for developing malaria early warning systems, particularly for Africa. For countries on which we have focused – Thailand, Indonesia and Afghanistan – there have been few studies to examine the dependency of malaria cases on these factors.

In an endemic area, the local adult population may acquire sufficient immunity after repeated infections. The disease could be deadly, however, to young children, pregnant women, those with depressed immunoresponse, and people new to the area. Because malaria is virtually nonexistent in the U.S., Americans traveling abroad and U.S. oversea forces are particularly vulnerable.

1.2 Why the military needs malaria modeling and early warning

The U.S. has considerable military forces overseas for normal and contingent operations. Currently, there are military installations in at least 18 countries. Most of these countries are well developed, such as Britain, Germany, Italy, Japan and Korea, etc., where living standards is high and public health service is excellent. But even in a fairly developed country, malaria may still be present. A noted example is the spill-over of malaria

epidemics from North Korea into South Korea across the De-Militarized Zone, around which many U.S. military bases are located.

When a country is malarious, it would affect the neighboring countries, because infected mosquitoes can fly across borders. In addition, cross-border movement of refugees or transient populations, who may go back and forth for economic reasons, may significantly increase the malaria incidence in the neighboring countries. The best example is Thailand, where all the endemic provinces are border provinces.

At the other extreme are the forces in contingent operations, which often take place in under developed or developing countries with malaria transmission. The 2002-2003 military operations in Cote d'Ivoire and Liberia, the first and the second Gulf Wars, and the ongoing Operation Enduring Freedom in Afghanistan and Horn of Africa are all in this category. In such cases, the forces must operate under the most austere environments. Because of the lack of sufficient public health support in many of these countries, there are unreported and under reported infectious diseases. In general, no detailed information regarding the diseases, pathogens, their prevalence, vectors and vector ecology are available.

The case mortality rate for US overseas forces due to malaria may be low compared with that for the civilians in Africa. But it takes precious manpower and resources away from the military to care for those who are infected and in recuperation — an issue especially critical in the battlefields. General MacArthur made a comment on malaria in the Second World War, "This will be a long war if for every division I have facing the enemy, I must count on a second division in hospital and a third division convalescing from this debilitating disease." This comment is still valid today as it was 70 years ago.

The concerned military organizations do issue guidelines and warnings to personnel going overseas. But because of the often scarcity of disease information, it is difficult to provide more precise advice, and therefore the troops may be over or under prepared. The guidelines and warnings do get updated when more disease related information can be collected by the deployed troops. But wars can have devastating effects on vector-borne infectious diseases: Public health services and vector control may be disrupted; housing for large population may be damaged; displaced populations are exposed to the environment; new larval habitats may be created through altered environment. Against this backdrop, the troops without any immunity to malaria need to move rapidly over large unknown territory. Therefore the disease threat may still be uncertain and may change quickly after the troops are deployed.

1.3 Purpose of the Benchmark Report

This report reviews the systems engineering development steps through which the MMS Project has undergone. These steps include Evaluation, Implementation, and Verification and Validation. The report then discusses the capabilities of the MMS Project as well as the typical capabilities of a state-of-the-art malaria early warning system, which GSAT essentially is. Finally, we discuss the improvements the MMS Project may bring to malaria risk assessments and control.

2. SUMMARY OF SYSTEMS ENGINEERING ACTIVITIES

2.1 The goal of the MMS Project

The objectives of the Malaria Modeling and Surveillance Project are: 1) to identify potential malaria vector larval habitats for targeted larval control, 2) to assess malaria risks based on past epidemiological data as well as climatic and environmental variables, and 3) to simulate spatio-temporal malaria transmission distribution and test transmission hypothesis. The value and benefits of the project are: 1) reduced likelihood of larvicide, insecticide, and antimalarial resistance, 2) reduced damage to the environment, 3) reduced morbidity and mortality for U.S. overseas forces, 4) improved public health for local populations.

2.2 Remote sensing data or results used in the project

Since rainfall provides vector breeding sites and prolongs vector life span by increasing humidity, precipitation or precipitation anomalies is the attribute most frequently used for predicting malaria epidemics. It has also been shown, however, that rainfall or the lack of it has a complex effect on malaria transmission for various parts of the world (Kovats *et al.*, 2003). For example, although moderate rainfall may promote malaria transmission, intense and prolonged rainfall may flush away larval habitats and thus reduce transmissions. Similarly, lack of rainfall does not always reduce larval populations. On the contrary, lack of rainfall may create new habitats, such as pools and puddles, in some regions and therefore increase larval population. In addition, droughts may be deleterious to predator populations or may cause human populations with no immunity to move to areas endemic with malaria (Kovats *et al.*, 2003). These factors may indirectly increase overall malaria transmissions. As rainfall or the lack of rainfall may influence malaria transmission in so many complex ways, a simplistic mathematical representation between rainfall and malaria transmission would be inadequate for modelling malaria transmission.

Another meteorological variable that is often used for predicting malaria transmission is temperature. Warmer temperature hastens larval and vector development and therefore increases the rate of vector production (Craig *et al.*, 1999). It also shortens the sporogonic cycle to allow vectors a longer period to transmit malaria. In addition, warmer air holds more moisture and therefore enhances mosquito survivorship.

Relative humidity is important for the survivorship of malaria vectors. While it is not a standard remote sensing data product, relative humidity can be computed from dew point and air temperatures, which are usually provided for from satellite instrument that measures atmosphere properties

Vegetation is often associated with vector breeding, feeding, and resting locations. A number of vegetation indices have been used in remote sensing and Earth science disciplines. The most widely used index is the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979). It is simply defined as the difference between the red and the near infrared bands normalized by twice the mean of these two bands. For green vegetation, the reflectance in the red band is low because of chlorophyll absorption, and

the reflectance in the near infra-red band is high because of the spongy mesophyll leave structure. The more vigorous and denser the vegetation is, therefore, the higher the NDVI becomes.

NDVI has also been used as a surrogate for rainfall estimate. It is an effective measure for arid or semi-arid region. For tropical regions where ample rainfall is normally received, vegetation index is a less sensitive measure for estimating rainfall. The mean vegetation index over a region reflects the degree of urbanization or lack of vegetation. In this sense, NDVI in a grid cell is used as an indicator for the mean level of vegetation present in the cell.

When atmospheric parameters for the future are needed, we use the forecast output from the NASA Seasonal and Interannual Prediction Project (NSIPP). After our malaria risk prediction model is trained, feeding estimated future environmental and meteorological data into the model will produce the estimated malaria incidence in the future. Naturally, any uncertainties or errors in the seasonal and Interannual forecasts will give rise to uncertainties or errors in the malaria incidence output.

The geophysical parameters and the instruments or models from which they are derived are shown in Table 1.

Table 1. Geophysical parameters and the instruments or models from which they are derived. X indicates that data products are available.

	Ground Cover	Vegetation Index	Surface Temperature	Rainfall	Humidity
ASTER	computed	computed		inferred	
AVHRR	computed	X		inferred	
MODIS	computed	X	X		computed
TRMM				X	
NSIPP			X	X	X
Ikonos	computed	computed		inferred	

2.3 The Integrated System Solution diagram

An Integrated System Solution Diagram illustrating the important elements of the MMS Project and how NASA data and results may enhance our partner's decision support capabilities concerning malaria is shown in Fig. 1.

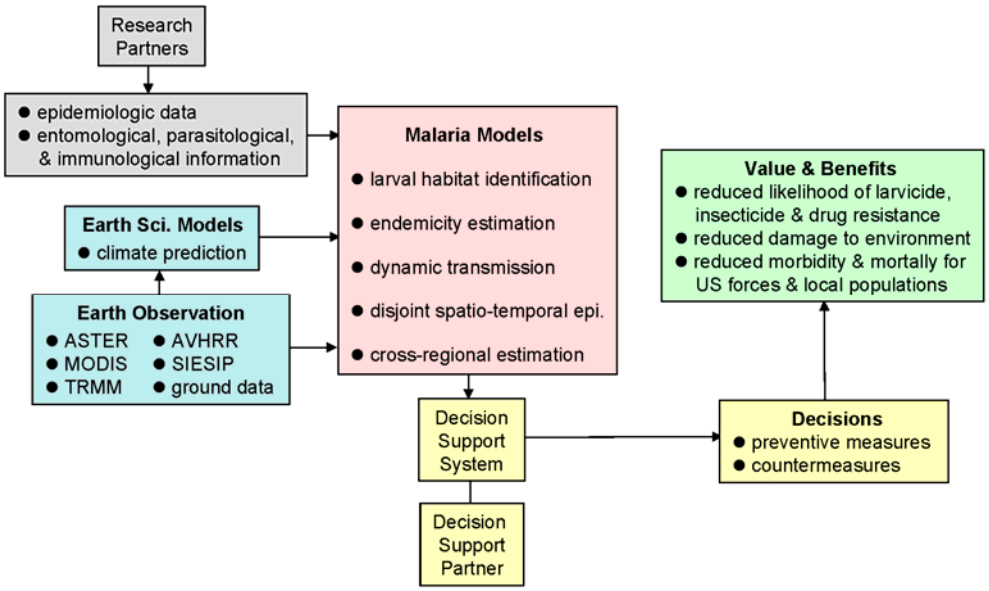

Figure 1. Integrated System Solution diagram. It illustrates the input and output flows, the malaria models in the MMS Project, and the value and benefits the project brings.

2.4 Summary of the Evaluation efforts

The Global Situational Awareness Tool (GSAT) is a system developed by the Air Force Special Operations Command (AFSOC) for assessing environmental and health issues of concerns for U.S. forces. Its predecessor is the Global Operational Environmental Review (GOER). GSAT incorporates human health and disease-related information in addition to environmental data. The system is built upon a set of linked databases with an intelligent, user-friendly interface. It is envisioned to be a knowledge- or rule-based system. The rules will be written by a group of experts to achieve outcomes as if human experts are performing similar analysis.

In the Evaluation Report, we concluded that the goals of GSAT and NASA's MMS Projects were clearly compatible, and that NASA data, results, and the output from MMS would enhance GSAT's capability. Output from the MMS Project, in particular the malaria risk assessments, will be input into GSAT. Certain NASA satellite data products are already included in GSAT's input streams. If needed, we will help GSAT to utilize other NASA data products, especially those used in the MMS Project.

When GSAT is fielded, the Air Force will gain a computerized environmental and medical planning capability. The combined capabilities of the malaria assessments and GSAT will provide the U.S. Air Force and its partners with a decision support tool valuable to U.S. military and civilian sectors. Because U.S. oversea forces generally assist the local public health organizations in disease prevention and control, the enhanced GSAT will also benefit the local populations.

2.5 Significance of the countries studied in the MMS Project

The countries for which we have studied the malaria risks include Thailand, Indonesia and Afghanistan. Each country has its unique malaria problems. Coincidentally, these are also the countries in which the U.S. has significant interest.

For example, Thailand is the main US ally in the Indo-China Peninsula. The Navy uses the port facilities in Thailand. And there are many US governmental and civilian organizations in Thailand. Every year, there is a large-scale, six-week joint Cobra Gold exercise with the Thai military and the participation from other Pacific allies. In the 2007 Cobra Gold exercise, five countries participated and 10 countries sent observers. Thailand is the most developed country in the Greater Mekong Subregion and has the best public health service in the region. The GMS itself, however, has the misfortune of being the world's epicenter of multi-drug resistant falciparum malaria. While Thailand has reduced its malaria burden in most provinces to a very low level, the malaria problem persists in the provinces next to Myanmar, Cambodia and Malaysia. Because of the displaced populations in these provinces and the transient workers crossing the border frequently to seek better economic opportunity, reducing malaria burden in the border provinces is difficult. The discovery of artemisinin-resistant falciparum malaria in 2008 in Cambodia is worrisome, as artemisinin is the most effective medicine and there is not new antimalarial on the horizon.

Indonesia is the largest muslin country in the world with a moderate political view and has a democratic government. Until several years ago, the country still grabbled with extremist and separatist movements and the remaining problems from the Asian economic crisis. However, political and economic situations have since much improved. Along with Singapore and Malaysia, Indonesia controls the 600 mile long Malacca Strait, through which approximately 35% of the world's oil production is transported on oil tankers. Threats to marine shipping from the modern-day pirates and occasionally from terrorists are problematic. The malaria endemicity was in decline in the early nineties. But most of the malaria control efforts were abandoned after the Asian economic crisis started. The government was also decentralized in this period and the responsibility for malaria control was relegated to the provincial government. Currently, the malaria endemicity on Java and Bali is low. But malaria is a problem on the outer islands. With over 18,000 islands, it is difficult to implement an effective malaria control strategies. In addition, topographic and geological characteristics in some regions, like the Menoreh Hills in Central Java, are difficult for effective larval control. Aside from drug-resistant falciparum malaria, there are serological evidences for the presence of drug-resistant vivax and malariae malaria. As prescription antimalarials are freely available, self-administered incomplete treatment is likely the major reason why the parasite species have developed drug resistance.

From the Soviet invasion in 1979 until today, there have been virtually non-stopped instability and military conflicts in Afghanistan. Currently there are ISAF and NATO forces from 37 countries working with the US in combat and other operations in various parts of Afghanistan. Because of years of instability, the public health infrastructure in Afghanistan had all but disappeared by 2004; the once successful malaria vertical control

program was long abandoned. And malaria has become a significant public health concern in Afghanistan despite its relatively cool climate and arid environment. Currently, approximately 60% of the population, or nearly 13 million people, live in endemic area. It is only since 2004, through the collaborations among WHO, Afghan Ministry of Public Health, NGOs and international donors, the efforts of rebuilding the public health infrastructure have begun. Nowadays, nearly 12,000 health facilities[1] have been established, along with clinics of various levels. It is the hope that the Basic Package of Health Services[2] (BPHS) can be afforded to a significant portion of the population. Although not a tradition practice, insecticide-treated bed nets (ITN) have been gradually accepted through extensive public awareness campaigns and mobile, subsidized sales. For some localities, more than 60% of the residents own bed nets. As indoor residual spraying (IRS) is difficult to implement in the current environment, broadening the ITN use is a viable approach for reducing malaria infections. The continued, consistent and proper use of ITNs, however, still requires sustained awareness campaigns.

2.6 Summary of the implementation efforts

To support the objectives of the MMS project, we have developed a number of mathematical techniques and models to assess malaria risks. These techniques are briefly explained in the following. Examples for using these techniques, as well as their quantitative performance measures, are given in Section 2.7, Summary of the V&V Efforts.

2.6.1 Textural-contextual classification methods

Ground cover classification using satellite imagery data of medium spatial resolution, such as Landsat data, normally involves only spectral information – that is, the radiance or instrument count in each spectral band. Using the same techniques on high spatial resolution data, such as Ikonos or Quickbird data, may lead to abysmal classification accuracy. This is most often due to the fact that shadowing effects – either self shadowing or shadows cast by nearby objects – become proportionally very important when instrument footprint is very small. To maintain classification accuracy, spatial information must also be used in addition to spectral information. Toward this end, we have developed techniques that utilize both textural and contextual information in classification. Such techniques greatly improve the ground cover classification accuracy when high spatial resolution satellite data is used.

2.6.2 Neural network methods

Various methods and techniques, including generalized linear models, regression trees, genetic algorithms and neural networks have been used in the MMS project to statistically model the relationship between malaria transmission and environmental or contextual determinants. Neural Network methods are what we most frequently used for modeling malaria risks.

---

[1] Including District Hospitals, Comprehensive Health Centers, Sub-Health Centers, Basic Health Centers, Mobile Clinics, and Health Posts.

[2] In addition to malaria, the BPHS consist of tuberculosis, HIV/AIDS, immunization, maternal and child health, Integrated Management of Childhood Illness (IMCI), etc.

This methods have been successfully used in many applications, including classification, regression, time series analysis, and handwritten character recognition (Nelson & Illingworth, 1990). In this approach, the probability density of the data is not assumed to follow any particular functional form. Rather, the characteristics of the probability density are determined entirely by the distribution in the data, hence, it is a data driven approach. This method is most suitable for problems that are too complex to be expressed in a closed, analytical form. For problems in which there are hidden, implicit variables, this approach is particularly suitable, as it is difficult to either specify the variables properly or sufficiently account for their effects mathematically. This method is called neural network because it resembles how biological neurons function (Gardner, 1993). Nodes in a neural network are analogous to neurons, the connections between the nodes are analogous to synapses. The behavior of the activation function corresponds to the firing of a neuron. The weights of the connections can be trained to give the aggregate of neurons a specific functionality. A network may accommodate complicated geometries in multidimensional space by incorporating hidden layers. Without hidden layers, the neural network method will be equivalent to the generalized linear model.

To train our neural network model, we feed observed or measured parameters from the past into the network. The input parameters may consist of meteorological, environmental and other variables; and the output is the corresponding malaria cases for that specific location and time. Once trained, the network will be able to estimate the cases at some other time period using the parameters corresponding to that time period.

2.6.3 Dynamic transmission model

Agent-based discrete event simulation is the technique we use for modeling the malaria transmission dynamically over time and space. In this model, the detailed interactions among the vector life cycle, the parasite's sporogonic cycle, and the human disease cycle are simulated for a real or realistic location. Extrinsic factors affecting malaria transmission are obtained through remote sensing or field observations. Intrinsic factors describing the parasitological and entomological characteristics are based on empirical or laboratory data. As parasitological, entomological and malariological knowledge is needed for such modeling, we collaborate with scientists at the Armed Forces Research Institute of Medical Sciences and the Walter Reed Army Institute of Research in developing this model. This model is useful for testing transmission hypotheses, such as larval and vector controls, abundance of breeding sites, housing construction, locations of farm animals, asymptomatic infections, availability of health care, arrival of refugees or foreign forces, etc. This model helps identify the most cost-effective ways for reducing malaria endemicity.

2.7 Summary of the V&V efforts

In this section, examples are given to illustrate how the techniques or models we have developed in the MMS project can be used for malaria risk detection, prediction and reduction. At the same time, the performance of these models can be obtained by comparing the modeled outcome with the actual condition. Thus these examples also serve the purpose of validating the models.

2.7.1 Risk detection

Common classification techniques can be used for identifying large larval habitats such as rice fields or water bodies. But special techniques must be used for identifying small larval habitats using satellite data with high spatial resolution. For example, *Anopheles sinensis* is the main vector species for transmitting vivax malaria in Korea. Irrigation and drainage ditches are the most common larval habitats for *An. sinensis.* In the following example, we use pan-sharpened Ikonos data, with an effective spatial resolution 1-4m, to identify these ditches. The per-pixel, spectral information only classification accuracy is approximately 72%. As shown in Fig. 2, various spatial techniques can be used to improve the classification. Using a textural-contextual technique, the accuracy can be enhanced to 92%. A classification result is shown in Fig. 3.

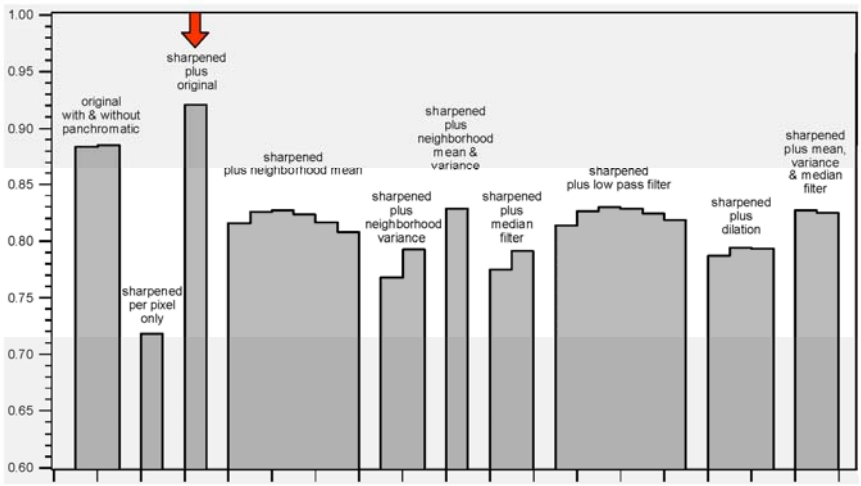

Figure 2. Classification accuracy of irrigation and drainages ditches for a test site in Korea using pan-sharpened Ikonos imagery data.

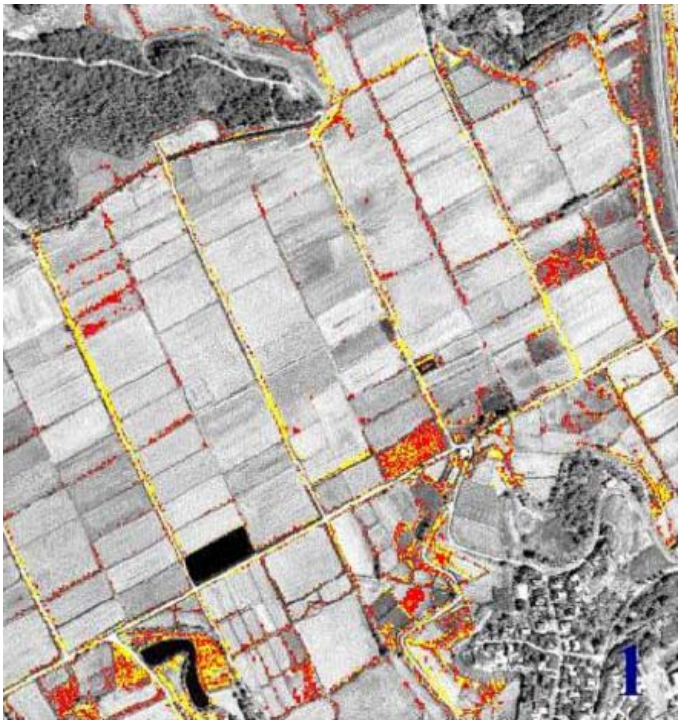

Figure 3. Some of the irritation and drainage ditches identified with the textural-contextual technique.

2.7.2 Risk prediction

Once a model is trained with past epidemiological data for a region, current and future malaria prevalence can be estimated by using current or future environmental parameters as input to the model. In the hindcasting (or retrospective forecasting) mode, the model is used to estimate the historical cases. The model's estimation accuracy can then be determined by comparing the model output with the events that took place in the past. Fig. 4 shows the comparison of actual provincial malaria prevalence in Thailand (left panel), and the hindcast prevalence (right panel). The hindcast result closely resembles the actual distribution, and is sufficiently accurate for all practical public health purposes. It can be observed that the hindcast slightly overestimates the malaria prevalence. This simply means that the malaria control effort implemented in the country was effective and result in a lower prevalence than what was warranted by environmental condition.

Actual, modeled and hindcast malaria prevalence can also be displayed together for a single province to show seasonal variation and how prevalence changes over time. For example, Fig. 5 shows such time series for the Takhar Province of Afghanistan. Takhar is one of the high risk provinces designated by WHO EMRO (Fig. 10). The training result (blue) follows the actual time series (red) closely. The hindcast malaria prevalence slightly under-estimates the actual prevalence, indicating the malaria situation is slightly worse than what is warranted by environmental factors. As Afghanistan has been in an unstable situation, this could be attributed to quite a few factors.

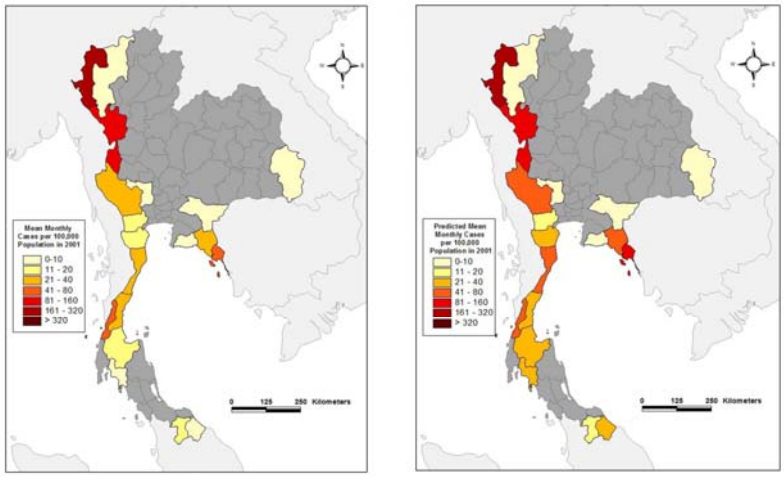

Figure 4. Actual malaria prevalence (left panel) and hindcast malaria prevalence (right panel) for the more endemic provinces in Thailand, from Kiang *et al.* (2006)

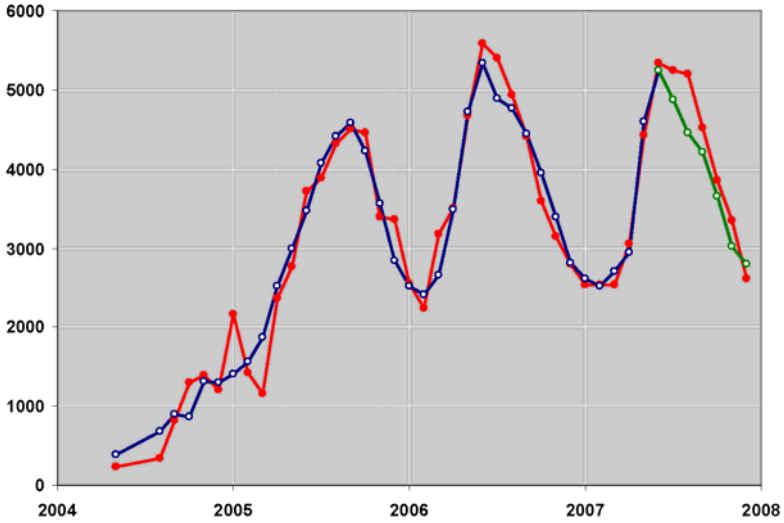

Figure 5. Actual (red), modeled (blue), and hindcast (green) malaria cases for the Takhar Province of Afghanistan.

2.7.3 Risk reduction

The dynamic transmission model simulates malaria transmission under various scenarios. As explained in Section 2.6.3, this model can be used to identify the key factors that perpetuate malaria transmission in a region. It can also be used to identify the most cost-effective measure to reduce malaria transmission. This model is validated by comparing the simulated results with the field data at a Thailand test site. AFRIMS and WRAIR have worked on this test site for more than five years, and may continue using it for testing malaria vaccines. Among other activities, the field work involved taking blood samples from most villagers every two weeks. Because it is very labor intensive and

rather costly to conduct field work, field measurements do not have the temporal resolution that is easily achievable in simulation. For example, simulated malaria prevalence at three different mosquito loading, as well as field measured prevalence with falciparum and vivax malaria combined is shown in Fig. 6. The middle, blue curve is for the simulated prevalence at a nominal mosquito loading. Because all residents tested positive were given antimalarials, prevalence decreases year after year.

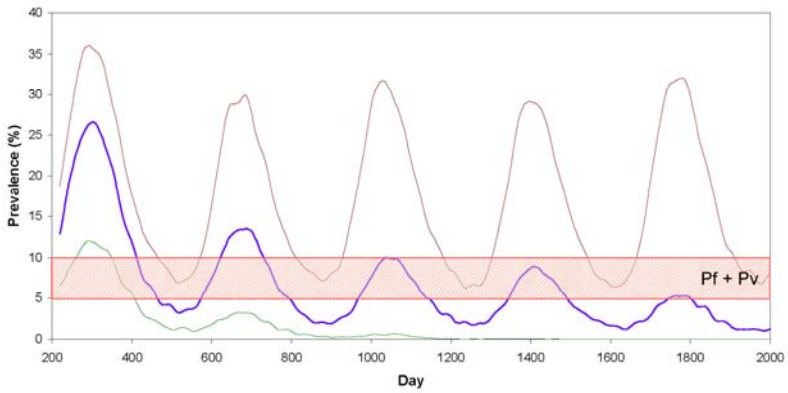

Figure 6. Simulated malaria prevalence for three mosquito loadings (curves) and field measured prevalence with falciparum and vivax malaria combined (band). Good agreement is shown.

Fig. 7 shows the simulated sporozoite rate, which is the fraction of all the mosquitoes that are infected with either falciparum or vivax malaria, at three mosquito loadings. The blue curve is at nominal mosquito loading. The range of *An. minimus* that were caught and shown infected with either the falciparum or the vivax parasites are shown in bands. Altogether, approximately 11,000 *An. minimus* were captured. Needless to say, our collaborators undertook tremendous efforts to analyze the field work results.

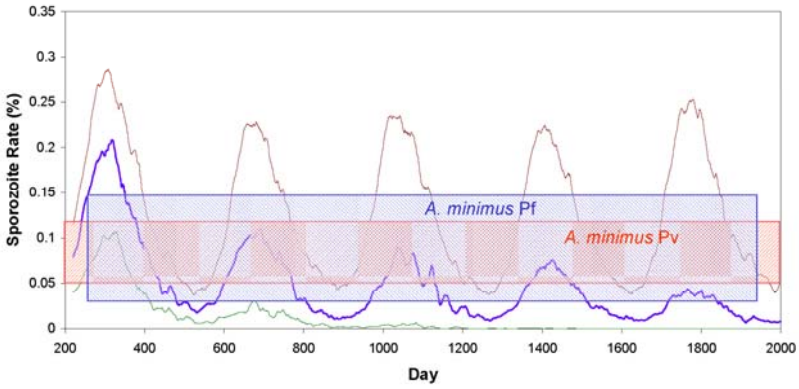

Figure 7. Simulated sporozoite rates and comparison with field observed sporozoite rates for *An. minimus* infected with *Plasmodium falciparum* or *vivax.* Good agreement is shown.

3. BENCHMARKING

In this Section, we will describe the representative, state-of-the-art capabilities for malaria risk assessments derived from remote sensing data. These capabilities were developed by university affiliated groups commissioned by the WHO or NGOs, by DoD contractors, or by government agencies to assess malaria risks. The GSAT is built upon these prevailing capabilities and also uses some of their output. Its capabilities for malaria risk assessments are essentially on a par with these modern systems.

Comparing them with the key capabilities developed in our MMS Project, as shown in the examples in Section 2.7, it will be apparent that the MMS capabilities offer more accurate and more precise risk assessments at a verifiable confidence level.

In addition to the state-of-the-art capabilities, the GSAT also consolidates other data and ground information, much of it classified, for its decision making. For example, additional layers of Bayesian analyses using other data available to GSAT can indeed improve the assessment accuracy. Such combination would effectively enhance the overall capabilities. Because the benchmark process has limited scope and involves only open-source materials, the conclusion given herein should not be construed as a reflection of GSAT's inadequacy.

The readers must also be reminded that the state-of-the-art capabilities of the contemporary systems, including our own MMS Project, continue to evolve. Therefore, the comparison and conclusion given in this section will only remain valid before any systems evolve significantly.

3.1 Quantitative dynamic vs. qualitative static risks

In modern, state-of-the-art malaria risk assessment systems, the risks are typically expressed in qualitative terms. A region can be designated as malaria risk free to high risk or to very high risk. If quantitative labels are associated with qualitative risk categories, the range for each category is usually very broad. In addition, such risk is often presented as a static risk, even though malaria risk is known changing with time and with season. In the following, we will use four examples – three from WHO and one from a NOAA project – for illustration.

3.1.1 World malaria risk in WHO 2005 report

The following example, as shown in Fig. 8, illustrates the world's malaria risk in 2003, as published by WHO in 2005 (WHO & UNICEF, 2005). Five qualitative categories of malaria endemicity– no malaria, low, moderate, high, and very high – are used without indication of number of malaria cases. This representation is also static, with no variation with respect to month or season.

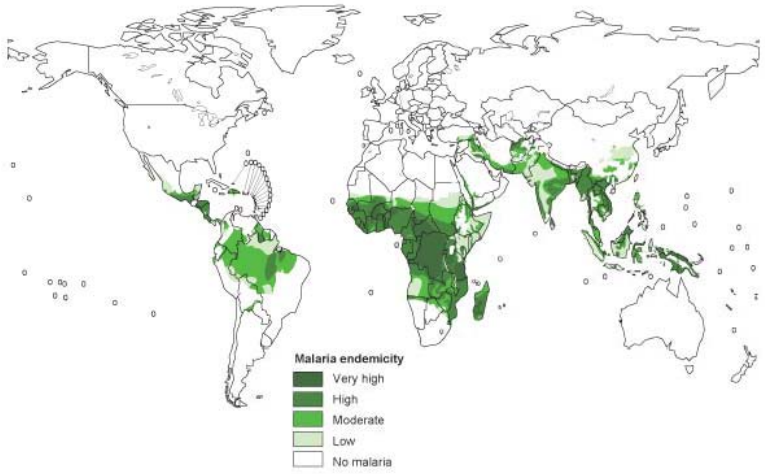

Figure 8. Global distribution of malaria endemicity in 2003, from WHO World Malaria Report 2005.

3.1.2 World malaria risk in WHO 2008 report

In this example, as shown in Fig. 9, five categories of malaria incidence are given to indicate the malaria incidence at country level (WHO, 2008). There is a quantitative scale given. Each category, however, is fairly broad.

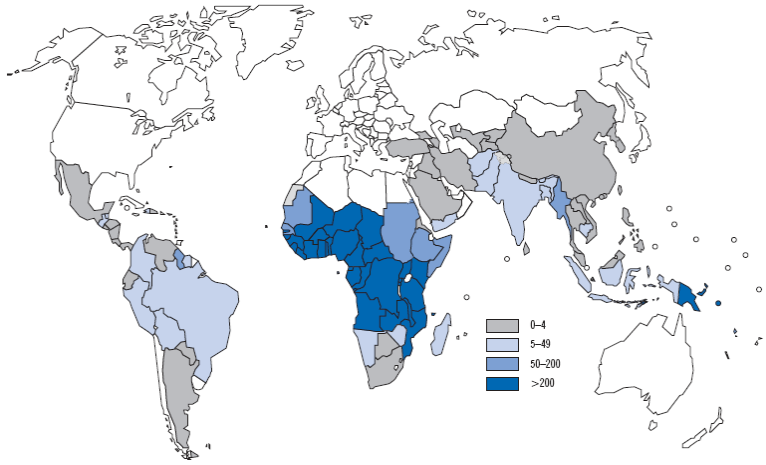

Figure 9. Estimated incidence of malaria per 1000 populations in 2006, from WHO World malaria report 2008.

3.1.3 Afghanistan malaria risk in WHO EMRO 2009 report

This example, as shown in Fig. 10, illustrates the malaria risk in Afghanistan as reported by WHO EMRO (2009). Three levels of static risks are shown.

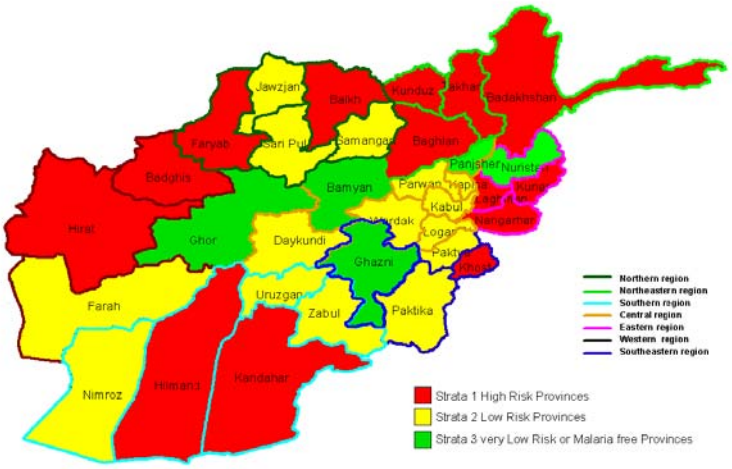

Figure 10. Malaria risk distribution at provincial level in Afghanistan as reported by WHO EMRO and by Afghanistan Ministry of Health.

3.1.4 Malaria risk map from a NOAA project

This example, as shown in Fig. 5, is based on using *vegetation health index* to monitor mosquito-borne diseases (Kogan, 2009). Three static qualitative categories of malaria risk – low, mixed, and high – are given.

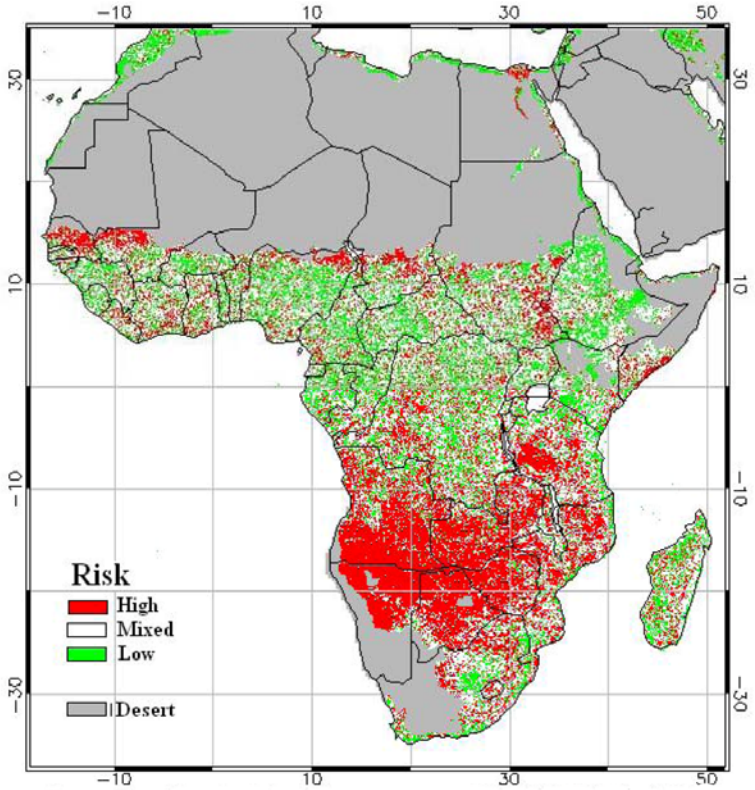

Figure 11. Malaria risk map derived from vegetation health indices.

3.2 Integrated vs. Climate suitability risks

Quite a few state-of-the-art malaria early warning systems, as well as some early warning systems for other diseases, utilize climate suitability as the basis for deriving disease risks. In the following, we will explain the limitation of such approach and the potential issues such approach may bring to disease risk assessments.

3.2.1 The meaning of climate suitability

The range of rainfall and temperature needed to maintain stable malaria transmission is called climate suitability (MARA/ARMA, 1998; Craig *et al.*, 1999; Small *et al.*, 2003; Omunbo *et al.*, 2004; Hay *et al.*, 2004). For example, in African regions where the *Anopheles gambiae* complex is the dominant species for transmitting falciparum malaria, climate suitability is associated with a temperature between 18º and 32º C and a rainfall exceeding 80 mm per month for at least 3 to 5 months (MARA/ARMA, 1998; Craig *et al.*, 1999).

Climate suitability is a useful concept for describing the potential for malaria transmission as long as the contextual determinants for malaria transmission, most of them remained implicit or unknown, do not vary significantly. A list of some of the possible contextual determinants is given in Fig. 15. In general, the concept of climate suitability is only applicable to a restricted spatial and temporal domain that has very similar characteristics to the domain where the climate suitability is obtained.

3.2.2 Transmission suitability for malaria in Afghanistan

Climate suitability can be expanded from rainfall and temperature to include other environmental parameters, such as elevation, distance to potential larval habitats, and other socioeconomic variables. Fig. 16 shows a more general suitability for malaria transmission in Afghanistan (Brooker *et al.*, 2006). The scale in the figure indicates the probability with or without transmission, with 1 indicating there is transmission, and 0 no transmission. This probability does not indicate transmission intensity or malaria prevalence. As with the classical climate suitability, care must be exercised in applying such criteria even in the same country. For example, housing construction, people's profession, personal protection, and socioeconomic status of the neighborhood may all override the suitability for transmission.

Figure 12. A more general suitability for malaria transmission in Afghanistan when parameters beyond rainfall and temperature are included, from Brooker *et al.* 2006). This probability does not represent transmission intensity.

3.2.3 Malaria risk map from Dynamic Technology

Fig. 13 shows a static malaria risk map derived from climate suitability for approximately 12 countries in Eastern Africa (Thomas *et al*., 2004). Five qualitative classes are given. Needless to say, the contextual determinants that affect malaria transmission in such a large geographic region may have significant variations.

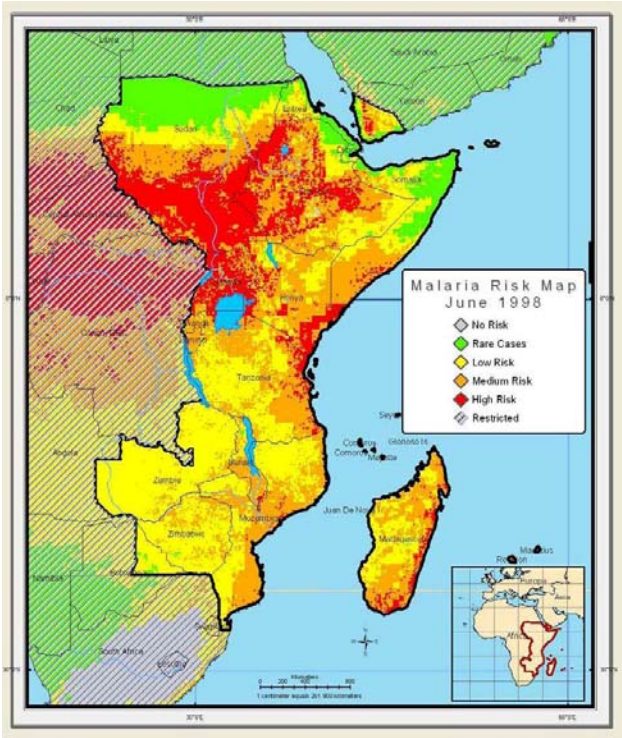

Figure 13. East African malaria risk map from Thomas et al. (2004).

3.2.4 Malaria risk map from MEWS Net

Fig. 14 shows a malaria risk distribution based on climate suitability for most of Africa. The same concern exists here, that contextual determinants, which also affect disease transmission besides environmental parameters, may vary widely across the continent.

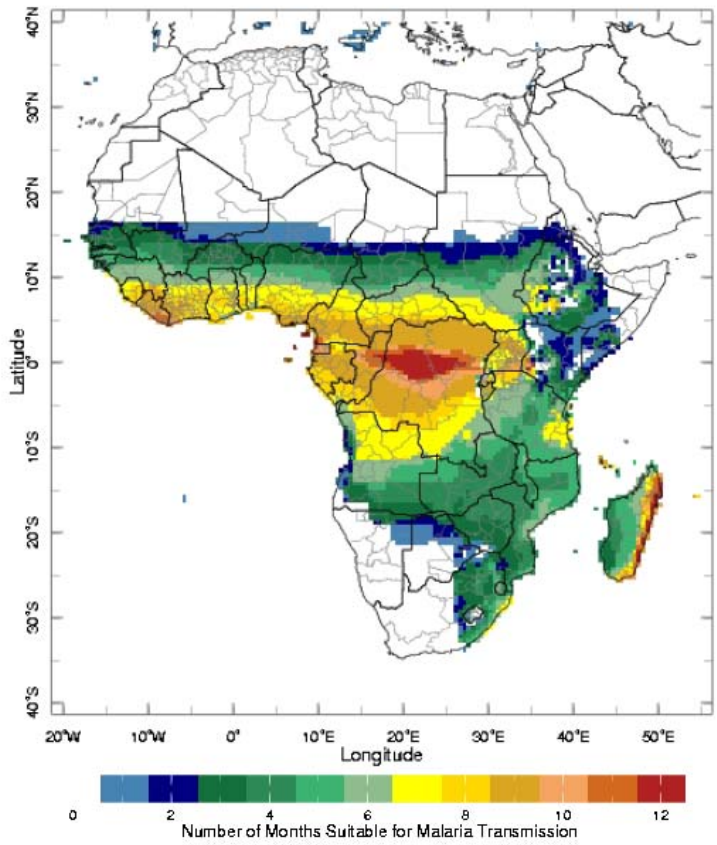

Figure 14. Climatological suitability for malaria transmission in 2009 from IRI website (ingrid.ldeo.columbia.edu/maproom/.Health/.Regional/.Africa/.Malaria/.CSMT)

3.2.5 What is missing in climate-suitability based reasoning

Climate suitability indicates how favorable the regional climate is for stable malaria transmission. Naturally, it depends on the ecology of the dominant malaria vector species, and varies with geographic region (WHO & UNICEF, 2005). Climate suitability can also be extended beyond the basic climatic promoters – rainfall and temperature – to include other environmental parameters, as shown in the Afghanistan example described previously in this section. In general, how much the malaria transmission potential associated with climate suitability can be materialized into actual malaria transmission depends on many other contextual determinants. Factors like socioeconomic conditions, public health infrastructures, herd immunity, irrigation and transportation projects, natural disasters, and military conflicts, have overriding effects on malaria transmission. Other contextual determinants very relevant to malaria transmission are shown in Fig. 15.

By the same token, merely using rainfall and temperature as the main determinants and extrapolate them to an entire region or continent to estimate malaria transmission is not scientifically correct.

To explain this problem using terminology in logic, climate suitability is a *necessary condition* for malaria transmission, but not a *sufficient condition*. A sufficient condition for malaria transmission would comprise all the contextual determinants that influence malaria transmission. As in general it is incorrect to treat a necessary condition as a sufficient condition, so is using climate suitability indiscriminatingly as the basis to estimate malaria risks.

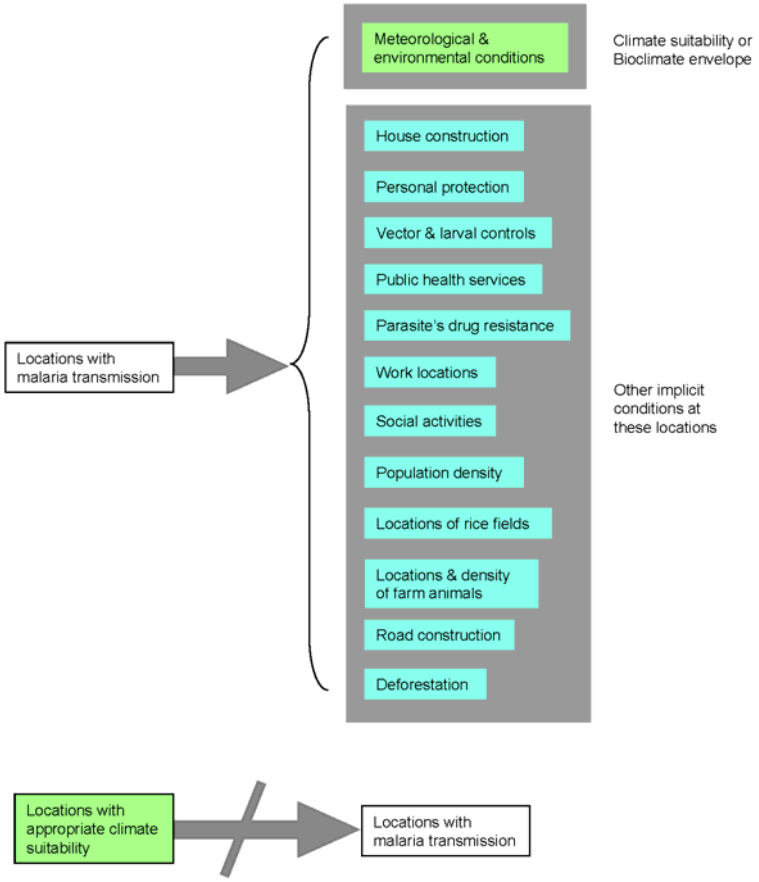

Figure 15. Climate suitability is a *necessary*, but not a *sufficient condition* for malaria transmission. Using only climate suitability to infer malaria transmission is logically incorrect and may lead to questionable conclusion.

Because it usually needs significant resources to gather the quantitative information on these background contextual determinants, many malaria studies or models simply leave them out. But simply ignoring them will run into the risk of treating necessary condition as sufficient condition. The way to avoid this dilemma is to use the actual local epidemiological data, as such data reflect the general constraints imposed by the

determinants that are not measured. In the MMS project, our risk assessments or prediction are always based on epidemiological data. Although it is often difficult to obtain epidemiological data, any such information, however little, would help to anchor a theoretical interpretation to reality. Our models would also allow disjoint spatiotemporal epidemiological data to be used (Fig. 1). In addition, by minimizing chi-squared errors instead of mean-squared errors, our models also have a built-in tolerance on the uncertainty of epidemiological records.

3.2.6 Why the US does not have malaria

Perhaps the best example for illustrating the fallacy of indiscriminately using climate suitability to derive malaria risk is the absence of malaria in the US. Malaria used to be a serious problem in North America. In the late 19th Century, the entire US to the east of the Rocky Mountains, except the Appalachian Mountains, was endemic with malaria (shown in Fig. 16), so were some regions at the West Coast and a significant portion of Canada. In 1946, the Communicable Disease Center was established by the Congress for the main mission of combating malaria. It was not until the 1950s when malaria was eradicated. And this agency became the predecessor of the current CDC. The climate in the US is still suitable for malaria transmission; and the malaria mosquito vectors are still present in the environment. But the mosquito vectors no longer carry the plasmodium parasites. Nowadays autochthonous transmission of malaria is very rare (CDC, 2004), and mostly can be traced to malaria infections acquired overseas. If the logic of climate suitability for malaria transmission were to apply to the US, most of US would be classified as endemic today.

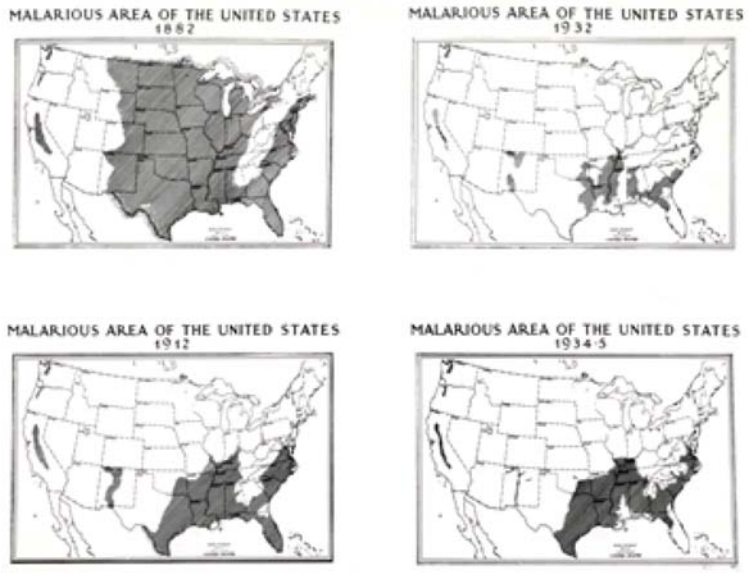

Figure 16. Historical malaria distribution in the US (from www.cdc.gov/malaria/history/index.htm).

.

3.3 Validated vs. hypothetical risks

State-of-the-art malaria risk assessment systems often do not have a built-in validation step. This may be simply due to the fact that traditionally qualitative system is not subject to quantitative analysis. Lack of validation, however, may leave a user wonder how much confidence and uncertainty can be associated with the risk assessments, especially when the risk is described qualitatively. On the other hand, all the results and outcomes from our MMS project are quantitative, and lend themselves to validation. Performance measures are associated with both the training and the prediction stages of our models. Hindcasting, or retrospective prediction, is what we normally use to derive model performance. In hindcasting, past environmental data are input into our trained models to estimate the past malaria cases. The performance can then be obtained by comparing the difference between the actual malaria cases with the modeled cases. As an example, Fig. 17 shows the malaria time series for Central Java, the trained malaria distribution, and the predicted distribution. Both the training and prediction performance can be easily computed and expressed as root-mean-squared (RMS) errors or other measures. Validating is therefore straightforward and a regular step in the modeling process.

On the other hand, when malaria risks are expressed in qualitative terms, such as risk-free, low, medium, or high risk, there may be no consensus on how the modeled risks can be validated, especially when there are no malaria case numbers associated with each risk categories.

4. CONCLUSIONS

The Global Situation Awareness Tool is a multi-purpose decision support system at the Air Force Special Operations Command. Initially developed as a civil engineering tool, it has gradually expanded to include other functionalities, including public health, disease risk assessments and prevention. GSAT consists of both classified and unclassified components. The unclassified component for disease risk assessment is partially based on some other systems or uses their output. Overall, GSAT's capabilities for disease risk assessment are on a par with the modern, state-of-the-art systems.

We use a number of contemporary systems to illustrate the characteristics of the state-of-the-art capabilities for disease risk assessments. Examples are drawn from the work of WHO-UNICEF RBM, WHO EMRO, Dynamic Technology, NOAA, University of London School of Tropical Medicine and Hygiene, and Columbia University IRI's MEWS Net. The output of these systems or projects can be characterized as – qualitative and stationary risk, climate suitability based risk, and hypothetical risk. Correspondingly, we show that the output from our Malaria Modeling and Surveillance Project can be characterized as – quantitative and dynamic risk, integrated disease risk, and validated risks. Our project can therefore provide more accurate, more precise risk assessments at a verifiable confidence level.

Two other capabilities in our MMS Project – textural-contextual classifier and agent-based dynamical malaria transmission model – are useful for detecting small larval habitats and identifying cost-effective ways to reduce malaria transmission. Similar capabilities are not available in these contemporary systems.

It should be noted, however, that all systems evolve, including our own. Conclusions from the comparison based on a snapshot of the current capabilities will only remain valid before any systems evolve substantially. Furthermore, GSAT's decision is based on a wide variety of information, both unclassified and classified. Such combination allows additional layers of analyses to be performed and may amplify its overall capability. Hence the conclusion given here should not be construed as a deficiency of GSAT.

ACRONYMS

AFRIMS	Armed Forces Research Institute of Medical Sciences
AFSOC	Air Force Special Operations Command
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BPHS	Basic Package of Health Services
CDC	Centers for Disease Control and Prevention
EMRO	Eastern Mediterranean Regional Office
GOER	Global Operational Environmental Review
GSAT	Global Situational Awareness Tool
IRI	International Research Institute for Climate and Society
IRS	indoor residual spraying
ISAF	International Security Assistance Force
ITN	insecticide-treated bed nets
MEWS	Malaria Early Warning System
MMS	Malaria Modeling and Surveillance
MODIS	Moderate Resolution Imaging Spectrometer
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NOAA	National Oceanic and Atmospheric Administration
NSIPP	NASA Seasonal and Interannual Prediction Project
RBM	Roll Back Malaria
RMS	root-mean-squared
SIESIP	Seasonal Interannual Earth Science Information Partner
SRTM	Shuttle Radar Topography Mission
TRMM	Tropical Rainfall Measuring Mission
UNICEF	United Nations Children's Fund
V&V	Verification & Validation
WHO	World Health Organization
WRAIR	Walter Reed Army Institute of Research

REFERENCES

Brooker et al., 2006. Spatial epidemiology of *Plasmodium vivax*, Afghanistan. Emerging Infectious Diseases, vol. 12, No. 10, 2006. www.cdc.gov/eid

CDC, 2004. Multifocal Autochthonous Transmission of Malaria Florida. MMWR Weekly, May 21, 2004 / 53(19) 412-413.

Craig MH, Snow RW, le Sueur D, 1999. A climate-based distribution model of malaria transmission in sub-Saharan Africa. Parasitology Today 15, 105-11.

Gardner D, eds., 1993. The neurobiology of neural networks, MIT Press, 227pp.

Hay SI, Guerra CA, Snow RW, 2004. Report on Agreement to Perform Work (APW): Determination of populations at malaria risk. Ref. M50/370/19, TALA Research Group, Department of Zoology, University of Oxford, 50pp.

Kiang R, Adimi F, Soika V, Nigro J, Singhasivanon P, Sirichaisinthop J, Leemingsawat S, Apiwathnasorn C, Looareesuwan S: Meteorological, environmental remote sensing and neural network analysis of the epidemiology of malaria transmission in Thailand. Geospatial Health 2006, 1:71-84.

Kogan F, 2009. Vegetation Health indices for Monitoring Mosquito Borne Diseases: Africa. Presentation slides.

Kovats RS, Bouma MJ, Hajat S, Worrall E, Haines A, 2003. El Niño and health. Lancet 362, 1481-89.

MARA/ARMA, 1998. Towards an Atlas of Malaria Risk in Africa – First Technical Report of the MARA/ARMA Collaboration. Mapping Malaria Risk in Africa/Atlas du Risque de la Malaria en Afrique, 45pp.

Nelson MM, Illingworth WT, 1990. A practical guide to neural nets, Addison-Wesley, 344pp.

Omumbo JA, Hay SI, Guerra CA, Snow RW, 2004. The relationship between the Plasmodium falciparum parasite ratio in childhood and climate estimates of malaria transmission in Kenya. Malaria Journal 3.

Small J, Goetz SJ, Hay SI, 2003. Climatic suitability for malaria transmission in Africa, 1911-1995. Proc Natl Acad Sci USA 100, 15341-5.

Thomas DM, Desch AD, Gaff HD, Scheele SK, Jordan RK, Davis JR, 2004. “Estimating infectious disease risk in the absence of incidence data,” ESRI International Health GIS Conference, Washington, D.C. October 17-20, 2004.

Tucker CJ, 1979. Red and photographic infrared linear combinations for monitoring

vegetation. Remote Sensing of Environment 8, 127-150.

WHO & UNICEF, 2005. World Malaria Report 2005. WHO/HTM/MAL/2005.1102, World Health Organization.

WHO, 2008. World Malaria Report 2008. WHO Library Cataloguing-in-Publication data WHO/HTM/GMP/2008.1.

WHO EMRO, 2009. Malaria country profile for Afghanistan. http://www.emro.who.int/rbm/CountryProfilesAFG