Doctoral Dissertation 2018

# Health Information System-of-Systems Management by Engineering Systems Multiple-Domain Modeling Approach Considering Spatiotemporal Dynamics

Suguru Okami

Supervisor

Professor Naohiko Kohtake

September 2018

Graduate School of System Design and Management,

Keio University

Major in System Design and Management

# SUMMARY OF DOCTORAL DISSERTATION

Title

Health Information System-of-Systems Management by Engineering Systems Multiple-Domain Modeling Approach Considering Spatiotemporal Dynamics

Abstract

Despite the substantial effort of various stakeholders, numerous people do not have sufficient access to healthcare services. Though the reliability and timeliness of health information collected in the healthcare system play critical roles in realizing appropriate healthcare resource allocation, the quality of health information continues to be of concern. Because the health information system is a system-of-systems that transforms over time in response to the changing environment and needs of stakeholders, it is important to understand the conditions of transitioning it continuously to improve the data quality.

This thesis aims to develop the continuous management cycle of the health information system through the iterative process of the mapping approach to estimate spatial heterogeneity of disease burden and its application for the modeling and analysis of the health information system for the purpose of better healthcare resource deployment while improving the quality of health information in the case of malaria issue in western Cambodia. The mapping approach with malaria spatial risk modeling is an effective tool and widely used for malaria containment actions, which was established through the long journey of effort by numerous contributors. However, as the disease burden of malaria decreases along the way malaria elimination effort progresses, this approach needs some adjustment in accordance with situational changes surrounding malaria epidemiology. The quality (accuracy) and reliability of data reported in the health information system continue to be of concern. The lack of quality health information is particularly alarming in situations with limited healthcare resources.

First, we applied the mathematical modeling approach to develop the spatiotemporal risk distribution model of malaria adjusted for the low-to-moderate malaria transmission setting by considering environmental context disparities surrounding human communities using routinely collected surveillance data, remote sensing data, and publicly available data to create the risk maps in fine-scale of two provinces (Pailin and Preah Vihear) in western Cambodia. The model was fitted to estimate the standardized morbidity ratio of malaria at each area. Fine-scale maps were created by the inverse distance weighed method and ordinal kriging interpolation of estimated risk. This approach was validated through comparing with the actual reported data from other sources.

Next, we developed an approach for modeling and analyzing the health information system-of-systems to understand the transitional complexities originated from the changes in the architecture, the process, and the surrounding environment of the system. We proposed the engineering systems multiple-domain modeling approach to model and analyze the transforming health information system-of-systems in the case of Cambodian malaria surveillance system. By using the attributes of the architecture, the process, and the environment within the model, relative weights of constituent systems were scored at each interval of time. This approach was validated through the comparison with the results of agent-based modeling simulation.

Finally, we demonstrated the transitions of the relative weights of constituent systems in the health information system-of-systems and its application for the adaptive resource allocation for a data quality intervention. By including the spatiotemporal dynamics of environmental attributes in the model along with the architectural and the process transformation, transitions of relative weights of constituent systems were scored at each interval of time. These score could be used to optimize the healthcare resource allocation continuously, whereas the considered architectural, process, and environmental changes could contribute to the sustained healthcare access by people.

Key Word (6 words)

System-of-systems; Health information system; Engineering systems; Spatiotemporal dynamics; Modeling; Complex systems.

## Table of contents

# List of figures

## List of tables

# ACKNOWLEDGEMENT

At first, I would like to express my deepest gratitude to professor Naohiko Kohtake, whose advice and suggestions as a first supervisor gave me the direction of my research. Furthermore, he provided me with various opportunities to study about geospatial science. Through my experience in the public health project of the G-SPASE program, I received not only a lot of instructive advices but also the support to conduct field study in Cambodia. This project became my starting point to find my research theme and provided numerous opportunities to connect with people. In addition, I would like to express my gratitude to professor Ryosuke Shibasaki of the university of Tokyo, professor Tomoki Nakaya of Tohoku university, and professor Hidekazu Nishimura of Keio university for reviewing my thesis as a secondary advisor and made a number of thoughtful suggestions for my research.

I would like to thank to professor Yoshiaki Ohkami and professor Dr. Taketoshi Hibiya from the SDM institute, graduate school of system design and management, Keio university for making a lot of advice and encouragement for my research. I would also like to thank all the laboratory members for sharing knowledge and discussions.

In addition, I would like to thank all the staff in Cambodia office of Foundation of International Development / Relief as collaborators of my research. Especially, Ms. Akemi Takahashi, Ms. Hiroko Oji, and Ms. Mao Sugita offered invaluable support and coordination of our field study since the time we visited Cambodia first time and of second visit for the field study and health practitioners interview.

This research was conducted as a part of the G-SPASE Program supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology and the Space Application Promotion Program funded by the NEC Corporation.

Finally, I would like to express my deepest gratitude to my wife, Kazuko Okami, for supporting me throughout years at SDM, Keio University. I can't thank her enough and would like to start to reciprocate her dedicated support from now on.

# 1. INTRODUCTION

## 1.1 Background

### 1.1.1 Problems of the access to healthcare

Despite substantial efforts by numerous stakeholders and advancements in health technologies, numerous people do not have sufficient access to healthcare services. According to the estimates by the World Health Organization (WHO), around 1.7 billion – approximately a third of the global population – did not have regular access to essential medicines and vaccines [1]. Furthermore, the situation has continued in the same range in recent years [2]. The issue of the access to healthcare services is a major problem and regarded as one of the prioritized global health agenda. In fact, the goal "Achieving universal health coverage, including financial risk protection, access to quality healthcare services, and access to safe, effective, quality, and affordable essential medicines and vaccines for all" has its place in the United Nation's sustainable development goal [3]. One of the major obstacles among wide ranging causes of the healthcare access inequality is the cost. There are substantial gaps of healthcare expenditure between low and high-income countries [4]. However, the cost is rather a part of complex web of factors. Various challenges such as the limited capacity of public health system and difficulties in distributing, prescribing, delivering, and using products still abound and require solutions in many places. The issue of healthcare resource gap is typically relevant in a remote places where the geographical access cannot be retained all time. It is also important to understand that the access to healthcare resources, e.g., a drug, vaccine, diagnostic, or other health product, does not automatically mean the improved health, especially in poor countries. Too often, patients are obtaining poor quality drugs that may cost money but have no impact on their health status. In many parts of the developing world, medicines are provided in bits of paper to the patients, with no instructions and no information for appropriate use. This practice can deleteriously affect quality and use of the health products and eventually the health outcomes of the patients [2]. Many stakeholders have made substantial efforts in a number of studies and projects to address the issue of healthcare access inequality. As a case, the geographical information system (GIS) is an effective tool for healthcare resource management. The methodology of spatial epidemiology is widely applied to understand the geospatial distributions of disease burdens and the level of health of people to reduce the risk of health [5]. In recent years, a number of technologies related to this research area, such as the people mobility analysis using mobile phone log [6], the application of remote sensing technology [7], big data analysis and simulation technologies to predict the effectiveness of interventions, have made significant progresses [8]. These technologies require the data and information collected in the healthcare system in any situations. Hence, reliable and timely health information plays critical roles for evidence-based decision-making of healthcare resource management.

### 1.1.2 Engineering challenges of the health information system

A health information system is the foundation that serves as a resource for such health information. Although the phrase "health information system" is not sufficiently clear, it was previously defined as "a set of components and procedures organized with the objective of generating information that will help improve healthcare-management decisions at all levels of the health system" [9]. A key component of health information system is public health serveillance, which helps in defining problems and providing timely basis of actions. While the quality and timeliness of health information are one of the topmost priorities, measuring health is conceptually and technically complex. Diverse stakeholders, not limited to the healthcare practitioners, are involved in the health information system at multiple levels. The requirements of these stakeholders may differ because of various factors, such as understanding of situations, positions, and viewpoints. Moreover, the requirements of the system may change over time in accordance with the change in the environment or new technological developments [10]. Hay et al. discussed such changes in the case of measuring malaria endemicity. When malaria becomes rare, along with the containment actions, detecting ongoing transmission becomes increasingly difficult using the commonly used parasite rate [11]. Hence, it is clear that the health information system needs to consider both the environmental and requirement changes continuously throughout its operational cycle. To maintain the quality data, various regular management actions, such as local quality control, up-to-date training, and frequent feedback to practitioners, are important [10]. Studies show that practical data-quality interventions improved the quality of data [12]. However, such corrective actions are often considerably expensive because of which they cannot be incorporated uniformly in the system. As a health information system is a combination of independent and interdependent systems, it can be classified as a system of systems (SoS) [13], which is transforming over time. This SoS can be characterized by the properties of the complex dynamic system such as diverse stakeholders and components, non linear behaviors, and capable of exhibiting emergence [14]. Because of the rapid technological development and growing complexity of the health environment and stakeholder needs, it becomes increasingly important to understand the condition of complex health information SoS in an ongoing manner for effective system engineering. This issue is particularly important in the situation such as the limited healthcare resource management in the complex healthcare system since the system serves as the supply source of data, which is used for the resource planning. So far, several studies have shown the quality of health information system [15-19]. The results were quite variable depending on the types of the tool used to collect the data and by countries [20]. Particulary, the lack of quality health information is alarming in situations with limited healthcare resources, which is observed in some less-individualized countries.

### 1.1.3 Malaria problems and emerging challenges toward elimination

Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected mosquitoes. For many years, malaria has remained an important global health threat that still results in hundreds of thousands of deaths every year [21]. Malaria is the 5$^{th}$ biggest cause of death in children except neonatal causes [22]. *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale* and *Prasmodium malariae* are the parasite species that can transmit to humans. Malaria damages the body through a number of pathways. Malaria causes the red blood cell distruction leading to anemia. The waves of parasites are bursting red blood cells, which can be a trigger for the classic cycles of fever and chills. The changes of adhesive properties of infected redblood cells can block the blood flow in the vessels causing tissue hypoxia. If this tissue hypoixia is happen in brain, it can cause cerebral malaria, which is often fatal. Once people get infected through the bites of vectors, infected *Anopheles* mosquitoes, they get acute febrile illness and several patients are in the severe condition such as hypoglycemia, anemia, respiratory distress and cerebral malaria, which can lead to the fatal conditions. Malaria also has chronic effects such as anemia, neurologic cognitive and developmental disfunction which lead to the impaired growth and development causes malnutrition as well as infant mortality and impaired productivity. However, the disease burden of malaria has significantly decreased in a number of malaria-endemic countries, due to substantial efforts made by many stakeholders. Now, decades after the global malaria eradication program, malaria elimination has begun to feature again on the global health agenda [23]. In recent years, an increasing number of countries such as Cambodia with low-to-moderate transmission areas have implemented actions to eliminate malaria from their entire territories [24]. In Cambodia, the target is to be malaria free by 2025 [25]. Recent activities have decreased the incidence of malaria in Cambodia to less than half the incidence in the years from 2000 to 2004 [26]. Approximately half the Cambodian population are living in malaria-free or in a low-transmission area [27]. However, a number of issues remain and new challenges are emerging in the effort to eliminate malaria. The emergence of artemisinin resistance, which has been reported mostly in the Greater-Mekong subregions, is one of the new challenges [28]. Artemisinin is a potent and rapidly acting blood schizonticide that is effective for all plasmodium species [29]. No alternative effective antimalarial treatment is available at present; therefore, the consequence could be dire if resistance spreads to wide geographical regions [30]. A number of reports have emerged of delayed parasite clearance in parasites in western Cambodia taking artemisinin [31-34]. In areas along the Cambodia–Thailand border, *P. falciparum* has become resistant and multi-drug resistance is a current major concern [31]. Usually in such areas, mobility of people is high, which contains the potential dangers of spreading the multi-drug resistance to larger geographical areas. One recent report showed that the artemisinin-resistant malaria parasite had the potential to infect vectors in other geographical regions [35]. The reported treatment failures in western Cambodia varied depending on the conditions [32, 36-38]; however, all the reports strongly emphasized the urgent need to address this issue. Appropriate medication is undoubtedly important and in areas such as those close to western Cambodia border, this approach occasionally needs intensive care and monitoring of patients. Several issues were reported in antimalarial drug use such as spreading availability of the artemisinin monotherapy, poor quality counterfeit medicines and unregulated antimalarial use [32, 39]. An important driver is said to be the use of oral artemisinins alone as monotherapy [30]. Along with the progresses of malaria containment activities, it becomes increasingly important to protect immunologically susceptible populations from serious consequences resulting from the reintroduction of malaria through residual foci. Migrations of asymptomatic patients have made it difficult to detect remaining transmission risk factors, and to protect people in malaria-free areas from the reintroduction of malaria [40]. To attain the desired outcomes, several studies such as those focused on screening and treatment [41], community-based surveillance [42], and mass drug administration [43] have been piloted. What is common to these interventions was the recognition that intensive support and engagement of local practitioners were critical in obtaining the desired outcomes. In addition, it is commonly observed that securing the same degree of investment for malaria containment as that obtained during the highly endemic period becomes more difficult along with the decline in malaria endemicity. Healthcare resources cannot be used inexhaustibly; therefore, identification of the target hotspots in malaria endemic areas, delivery of sufficient stockpiles of resources, and intimate support for local healthcare providers are essential, especially in remote endemic regions where accessibility cannot be retained over whole year.

### 1.1.4 Data quality issues in malaria surveillance

While the surveillance system is the foundation that serves as the data supply source for healthcare resource management, it has been reported that the quality (accuracy) and reliability of the data collected at health facilities continues to be of concern. One of the possible reasons for this data discrepancy is an inflation of reported data at the facility level to show the attainment of local targets. Similarly, inflated data on the facility report were found to occur, largely at health facilities with fewer financial resources and supervisory visits [44-46]. Moreover, previous studies of intensive focused screening have shown that many malaria cases were asymptomatic, which made it difficult to identify malaria cases effectively using conventional passive surveillance systems [11, 47]. Given that the health facilities are playing important roles in surveillance data collection, identification of malaria hotspots and the provision of appropriate supports to health facilities are important for maintaining the quality of surveillance data.

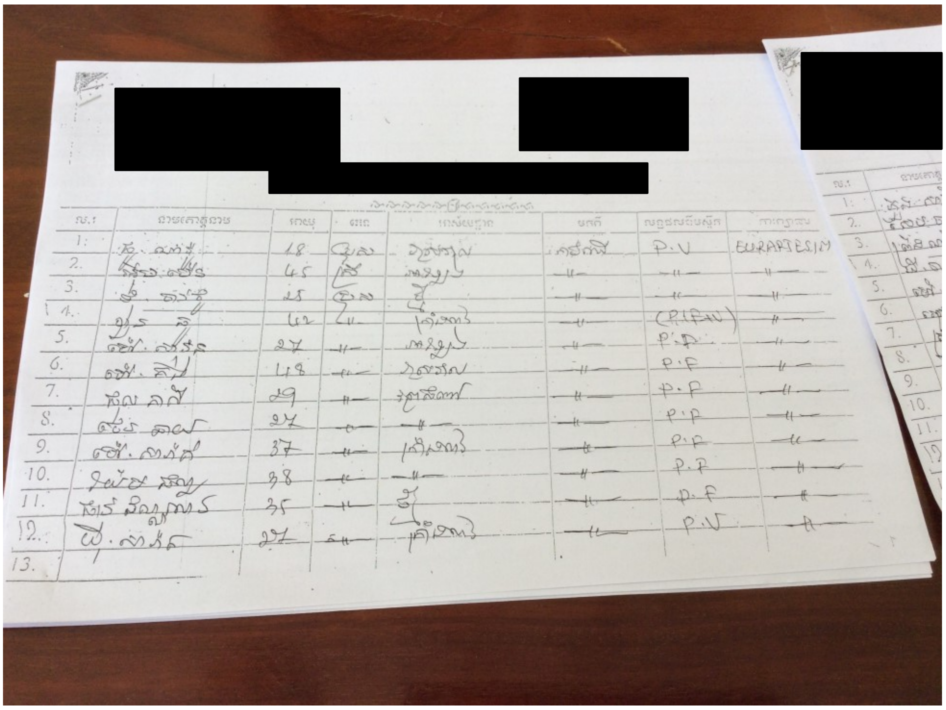


**Figure 1 Example of malaria case reporting form**

Figure 1 shows the example of the malaria case reporting form at a health center observed during the field study in a rural area of Kampong Chhnang province in Cambodia. At the time of the field study in February 2015, malaria cases were reported in the paper-based form without quality control measures such as the logical data quality check typically devised in the web-based case reporting system. The interview results with the director of health center indicate the resource inequality between health center staff and village malaria workers (Table 1). This case illustrates the challenges of reliable data collection in the health information system. If we can measure the risk and present it on the map, for areas not covered by current healthcare system sufficiently, the required amount of healthcare resource delivery and staff deployment can be facilitated. As healthcare resources cannot be used inexhaustibly, identification of the target hotspot of malaria endemic area, delivery of the sufficient stockpile of resources and intimate support for field healthcare providers are essential. However, regardless of these solutions, the quality (accuracy) and reliability of data reported in the health information system, i.e., the data source to develop GIS, is continued to be of concern.

**Table 1 Questions and responses at the interview with the local health center staff**

| No | Questions | Responses |
|---|---|---|
| 1 | How many patients have come to this center? (Today/ This week/ This month) | 50 / 350 / 1,500 |
| 2 | What kind of disease do patients have most? | Fever and Gastroenteritis |
| 3 | How about malaria patients, how much is the proportion? | Not a serious issue |
| 4 | Can you spare enough time for each patient? | Yes |
| 5 | Once patients are treated at this hospital do they come back here to check to ensure they are fully recovered? | Yes |
| 6 | For how long do you prescribe antimalarials – is it 3 Day, 6 Day, or 1 week? | 3 days |
| 7 | Are health resources and staffs enough to cover all the patients? | Yes |
| 8 | What kind of information do you need to provide enough care? | Medical consultation |
| 9 | Are you regularly cooperating with village malaria worker? | Yes |
| 10 | Do they have enough knowledge, experience and skills for treating patients appropriately? | Not enough |
| 11 | What is the issue in cooperating with village malaria worker? | Need more training and budget |
| 12 | How many patients are going to private health care facilities such as private pharmacy or clinics? | Almost 50% |
| 13 | Do you need more budget for providing quality care for patients? | Yes |
| 14 | Do you think patients will adhere to the treatment more and keep taking medications up to prescribed terms if they are followed up more intimately? | Yes |
| 15 | Do you have any other things to worry about? | I worry about some people move their home to other living and education is low. |

This issue is particularly important in the situation such as the limited healthcare resource management in the complex healthcare system since the system serves as the supply source of data, which is used for the resource planning. Furthermore, Cambodian malaria surveillance system was reformed its architecture and process to fulfill the changing stakeholders' requirements and transitioning environmental conditions, which makes it difficult to capture the overall conditions of the system. Corrective actions such as data quality intervention are often considerably expensive because of which they cannot be incorporated uniformly in the system. Hence, it becomes increasingly important to understand the condition of complex health information SoS in an ongoing manner. The strengthening of surveillance by improving the quality of data in the health information system, together with improving the treatment of infection, leads to a more sustainable effort to eliminate malaria.

## 1.2 Research objective

This thesis aims to develop the continuous management cycle of health information system through the iterative process of the mapping approach to estimate the spatial heterogeneity of disease burdens and its application for the modeling and analysis of the health information system for the purpose of better healthcare resource deployment while improving the quality of health information in the case of the malaria issue in western Cambodia. A schematic figure of the proposed system is shown in figure 2. The mapping approach with malaria spatial risk distribution modeling is an effective tool and widely used for malaria containment actions, which was established through the long journey of effort by many contributors. However, as the disease burden of malaria decreases along the way malaria elimination effort progresses, this approach needs some adjustment in accordance with the situational changes in the malaria epidemiology. The quality (accuracy) and reliability of data reported in the health information system continue to be of concern. The lack of quality health information is particularly alarming in situations with limited healthcare resources. The health information system is transforming over time while considering both the environmental and requirement changes continuously throughout its operational cycle. Hence, it is clearly important to understand the condition of complex health information SoS in an ongoing manner for effective healthcare resource management.

The specific objectives of this study were as follows:

(1) *To develop the spatiotemporal malaria risk model of malaria adjusted for the low-to-moderate malaria transmission settings, considering environmental context disparities surrounding human communities using routine surveillance data in the health information system, remote sensing data, and publicly available data.*

(2) *To develop an approach for modeling and analyzing health information SoS to understand the transitional complexities originated from the changes in the*

*architecture, the process, and the surrounding environment for effective system management.*

(3) *To demonstrate the transitions of the relative weights of the constituent systems in the health information SoS and its applicaiton for the adaptive resource allocation for data quality interventions.*

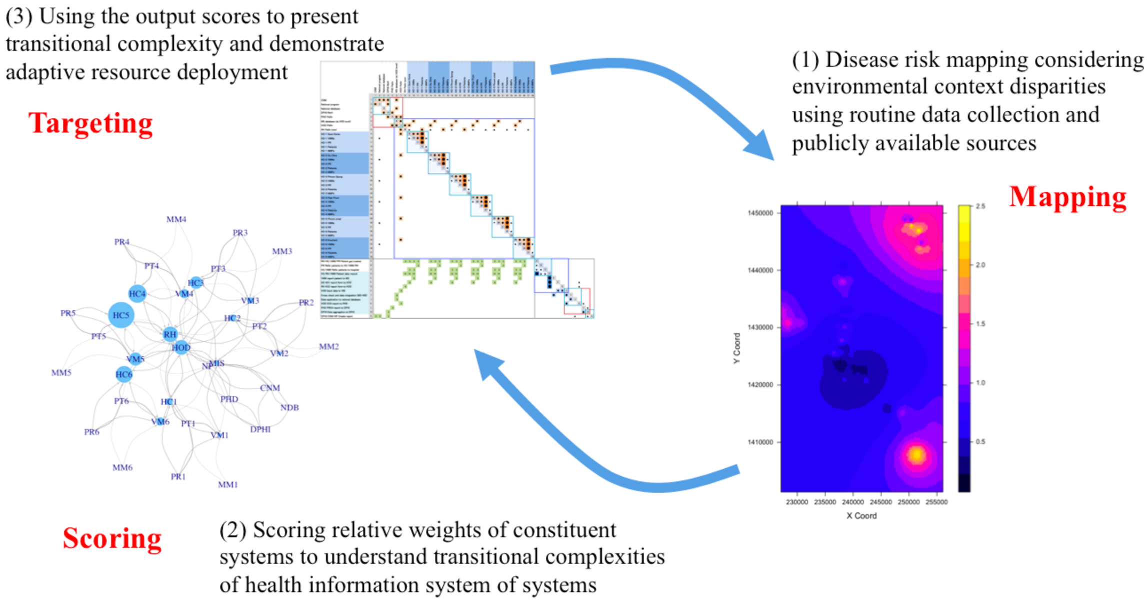



**Figure 2 Schematic figure of proposed system**

## 1.3 Contributions of this work

My original contributions are; (1) the demonstration of the mathematical modeling approach of spatial risk distribution of malaria adjusted for the low-to-moderate transmission setting from routinely aggregated surveillance reports considering the environmental context disparities surrounding human communities with low operating cost, (2) the development of the practical approach for the modeling and the analysis for continuous management of the transforming health information SoS, and (3) the demonstration of the transitional complexities of the health information SoS by scoring relative weights of constituent systems and its application for adaptive optimization of the health resource deployment for data quality interventions in the system.

Figure 3 presents the typical program phase of malaria elimination. The low-to-moderate transmission setting is the situation where the many of the countries aiming for malaria elimination will soon be facing and thus an important step toward the goal. At this stage, the program reorientation is needed for moving the program effort forward. The approach and implications described here will provide the method for more efficient resource allocation applicable not only in operating phase but also in design phase with low operating cost for countries under such situations.

This study covers not only the spatial risk distribution modeling of malaria but also its application in the health information management. The quality of data collected through the routine health information system continues to be of concern. Hence, there is a clear need for the practical approach for the effective health information SoS management. To the best knowledge of the author, this is the first study to provide the practical solution for addressing such questions by taking changes in the architectural, the process, and the environmental aspects of the SoS simultaneously and continuously in the sustainable cycle of the system. This study may also provide the useful insight for the future research and potential expansion to the health information SoS engineering in the other domains.

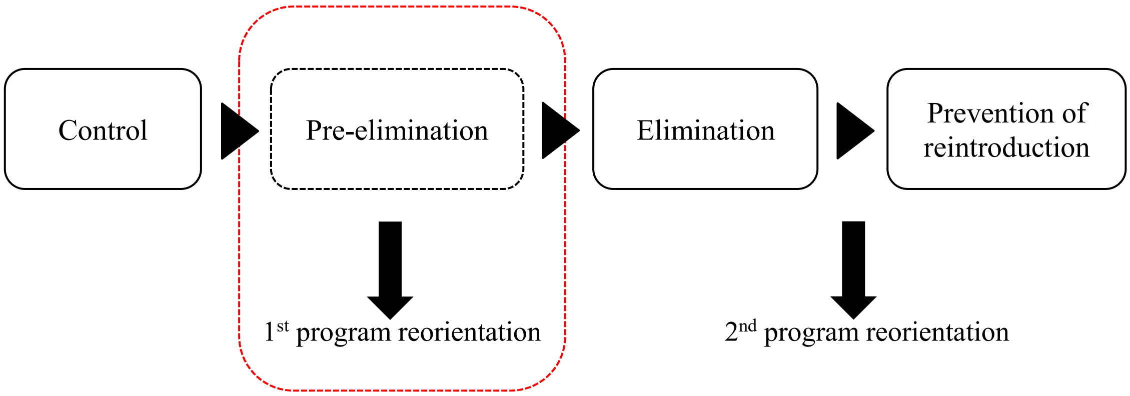



**Figure 3 Program phase in the malaria elimination program (created from [48])**

## 1.4 Chapter organization

Figure 4 shows the outline of this thesis that consists of following chapter organizations:

**Chapter 2** reviews previous works in related areas that include (i) Studies for improving the access to healthcare services, (ii) Mapping approaches of malaria disease burdens with spatial risk distribution modeling, and (iii) Studies of quality improvement and management of health related data in the health information system. Furthermore, the author explains the limitations of the current approaches and clarifies the requirement to address these issues.

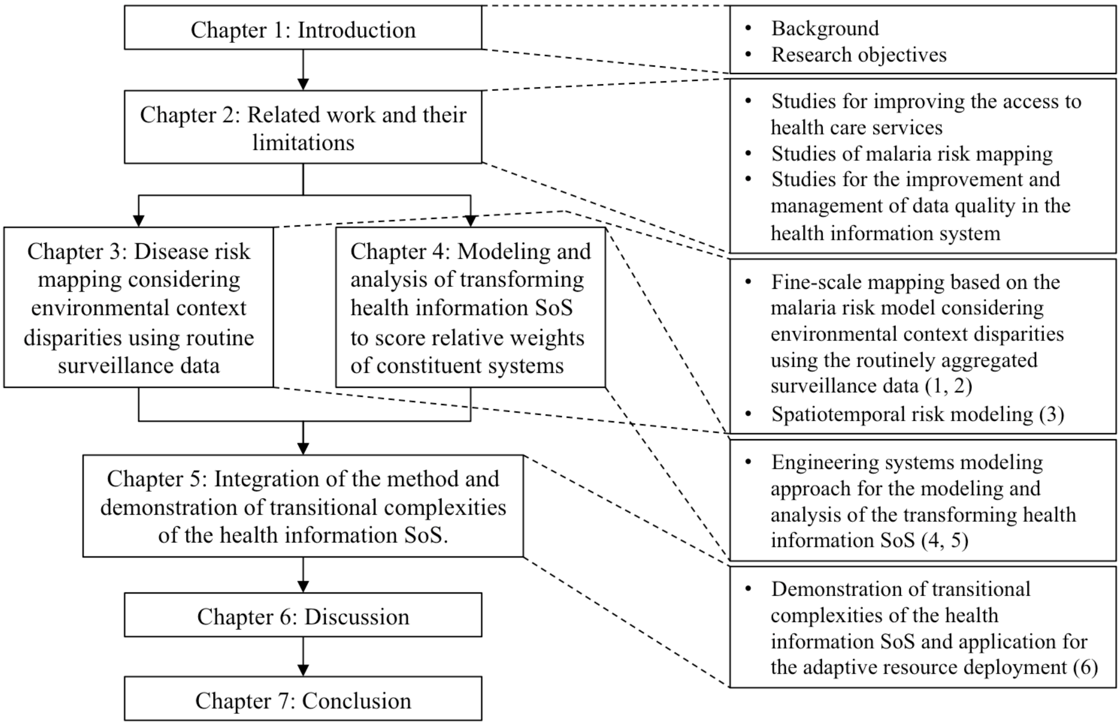
**Chapter 3** presents the development and evaluation of the spatiotemporal malaria risk model using the hierarchical Bayesian modeling frame corresponding to the environmental context disparities. The author presents the fine-scale mapping with the model developed. This approach was evaluated through the comparison with previously reported maps and geocoded data.

**Chapter 4** presents the demonstrations of the modeling and analysis approach, engineering systems multiple-domain modeling approach, of transforming health information SoS in the case of Cambodian malaria surveillance system. Author proposes a method to calculate the relative weights of constituent systems to understand transitional complexities of the health information SoS. This approach was validated through the computational simulation and the author discusses further insights obtained from the results.

**Chapter 5** consolidates the results presented in chapter 3 and 4. The author integrates the method explained in the previous chapters to build a continuous cycle of system and presents the transitional complexity of the health information SoS using the output scores of the system. The author also demonstrates the adaptive resource allocation of healthcare resources as a case of application.

**Chapter 6** discusses the results and findings of this study. Author provides the insights obtained throughout the development and evaluation processes of the proposed system for future work and extensive applications of the proposed system.

**Chapter 7** concludes this thesis. This chapter summarizes contributions of this study and describes the future study directions.




1. **Okami S**, Kohtake N (2015) Designing the GIS Predicting Regional Malaria Endemicity in Cambodia. Proc. of Esri Health and Human Services Conference 2015, Atlanta, GA.

2. **Okami S**, Kohtake N (2016) Fine-Scale Mapping by Spatial Risk Distribution Modeling for Regional Malaria Endemicity and Its Implications under the Low-to-Moderate Transmission Setting in Western Cambodia. PLoS ONE 11(7): e0158737. doi:10.1371/journal.pone.0158737

3. **Okami S**, Kohtake N (2017) Spatiotemporal Modeling for Fine-Scale Maps of Regional Malaria Endemicity and Its Implications for Transitional Complexities in a Routine Surveillance Network in Western Cambodia. Frontiers in Public Health 5: 262. doi: 10.3389/fpubh.2017.00262

4. **Okami S**, Kohtake N (2017) Modeling and Analysis of Healthcare Information System-of Systems for Managing Transitional Complexities Using Engineering Systems Multiple-Domain Matrix, Proc. IEEE International Systems Conference. p.688 – 695.

5. **Okami S**, Kohtake N (2017) Transitional Complexity of Health Information System of Systems: Managing by the Engineering System Multiple-Domain Modeling Approach. IEEE Systems Journal. doi: 10.1109/JSYST.2017.2778418

6. **Okami S**, Kohtake N (2018) Managing Health Information System-of-Systems by Engineering Systems Multiple-Domain Modeling Approach Considering Spatiotemporal Dynamics. IEEE International Symposium on Systems Engineering. (Accepted).

**Figure 4 Chapter organization**

# 2. RELATED WORKS AND EXISTING ISSUES

## 2.1 Studies for the improvement of the access to healthcare services

Since the healthcare access issue is one of the important global health agenda, many attempts to improve the access have been made in this area. However, only few of these efforts have been systematically and comprehensively documented and analyzed. Aday and Andersen, for example, developed a framework to study access to medical care in the United States in the 1970s [49]. In their framework, the role of health systems and population factors were addressed in shaping access to healthcare. A group at the London School of Hygiene and Tropical Medicine conducted another study of healthcare access and published a series of paper on “expanding access to priority health interventions” as the basis for analyzing the constraints to scaling up of healthcare technologies [50]. Frost and Reich conducted a study to develop a comprehensive analytical framework for access by examining six case studies [2]. They explicitly focused on health technologies by taking the six cases of specific health technologies and followed the flow of technologies through different phases of access (Figure 5) while incorporating global factors to develop the analytical framework. Figure 6 shows the access framework proposed by the research team. This analytical framework includes many processes involved in access to health technologies and is based on four A’s: architecture, the organizational structure and relationships for access; availability, which emphasizes the supply components of access; affordability, the cost issues for various payers; and adoption, which includes demand factors and acceptance. Each case study chosen for this study was summarized into this framework.

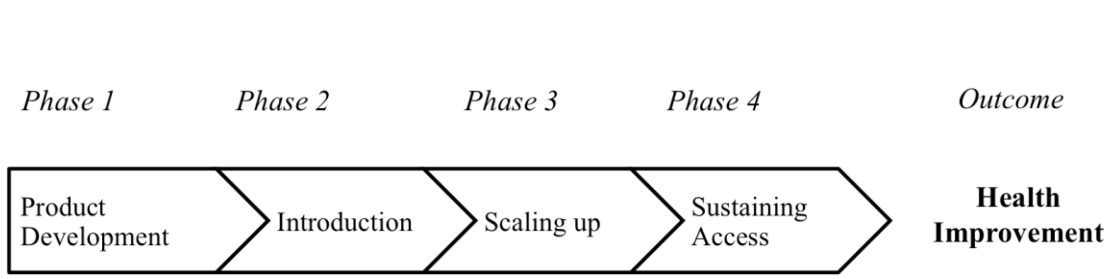



**Figure 5 The access phases of healthcare (created from [2])**

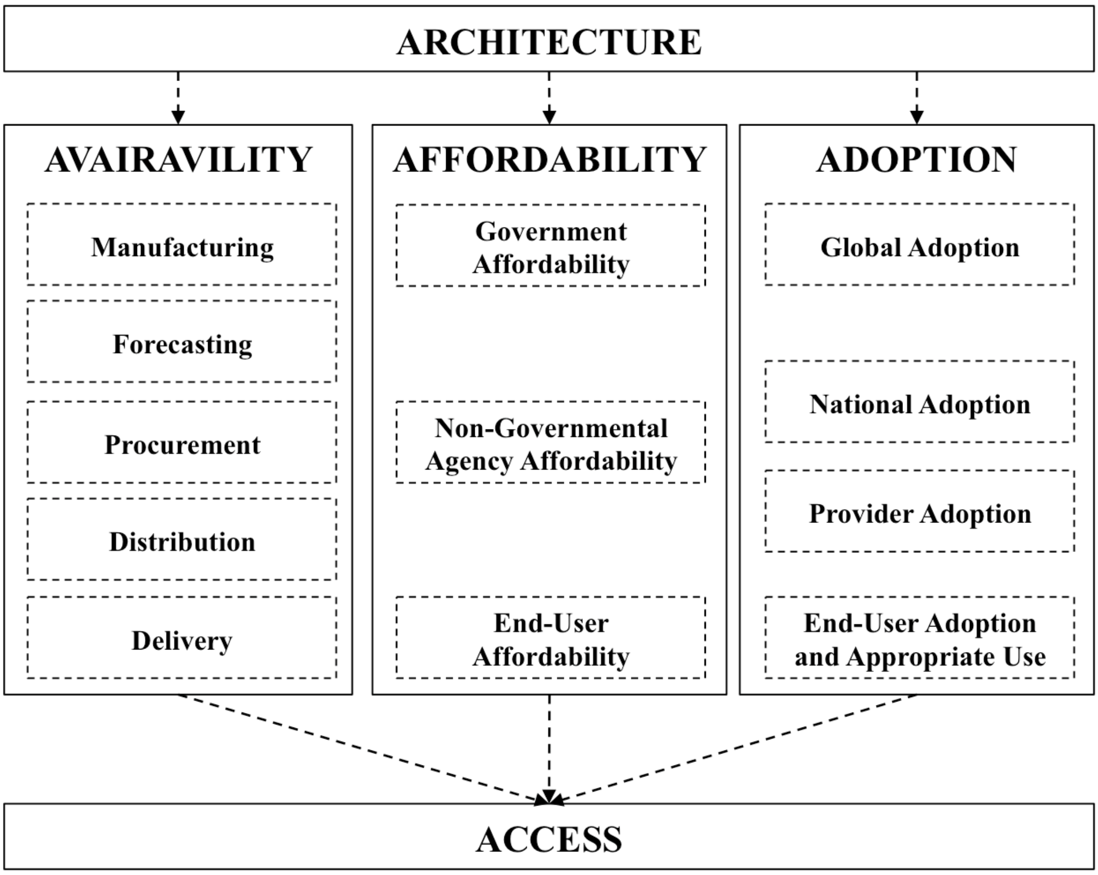



**Figure 6 The access framework (created from [2])**

Through the case studies, the complexity of creating access for health technologies in developing countries was demonstrated. Problems to access occur at many points along the pathway to the end-user. They synthesized the insights from the analysis of case studies into following 6 findings: 1, Developing a safe and efficacious technology is necessary but not sufficient for ensuring technology access and health improvement. Products do not fly off the shelf on their own; 2, Creating access depends on effective product advocacy. The case studies showed that product advocacy has three important components: a product champion, a coordinating architecture, and an access plan with strategies; 3, Access requires the creation and shaping of product adoption by four key groups: global experts, national policy maker, providers, and end-users; 4, The cost of health technology are needed to help expand access must address affordability; 5, Supply-side strategies that assure the availability of a technology are needed to help expand access for health technologies in developing countries; 6, Limited health infrastructure in many developing countries impedes technology access. Efforts to scale up access to technologies need to invest in health system strengthening to ensure sustained access. However, the case study chosen for their study did not examine the final access phase, where the individual countries seek to sustain the use of a technology for long-term prevention, control, or eradication of related disease. The aim in this thesis is to fulfill this space by addressing the issue of sustaining availability of healthcare services through the continuous cycle of improving healthcare resource deployment and health information system strengthening.

## 2.2 Mapping approaches of malaria disease burdens with spatial risk distribution modeling

Recent efforts to quantify the risk burden and the creation of spatial prediction map of malaria risk made substantial contributions toward identifying target hotspots [51-52]. Figure 7 shows the framework for geospatial science applied to malaria elimination [52]. Various kinds of data can be used such as the intervention coverage and infrastructure and target residences in view of operations. Also the remote sensing data, meteorological data from environmental side as well as malaria surveillance, survey and entomogical data can be used for the malaria quantifications. By integrating these data, GIS can be developed and the spatial statistical analysis can be conducted so that these data can be used for malaria containment activities as outputs. By predicting areas at risk and examining the effectiveness of interventions from estimated risk at the target hotspots, further steps closer to the efficient resource allocation can be attained. In this context, a world map of *P. falciparum* malaria endemicity was published using parasite rate (PR) surveillance report and the model-based geostatistical approach (Figure 8) [53-54]. These procedures were implemented within a Bayesian statistical framework to represent the uncertainty in the unknown map while retaining robustness of these predictions [55]. These mapping products and methodologies for spatial risk distribution modeling are provided by Malaria Atlas Project, which is aiming to disseminate free, accurate and up-to-date information on malaria. The team in the university of Oxford is receiving designation as a WHO Collaborating Centre in geospatial disease modeling.

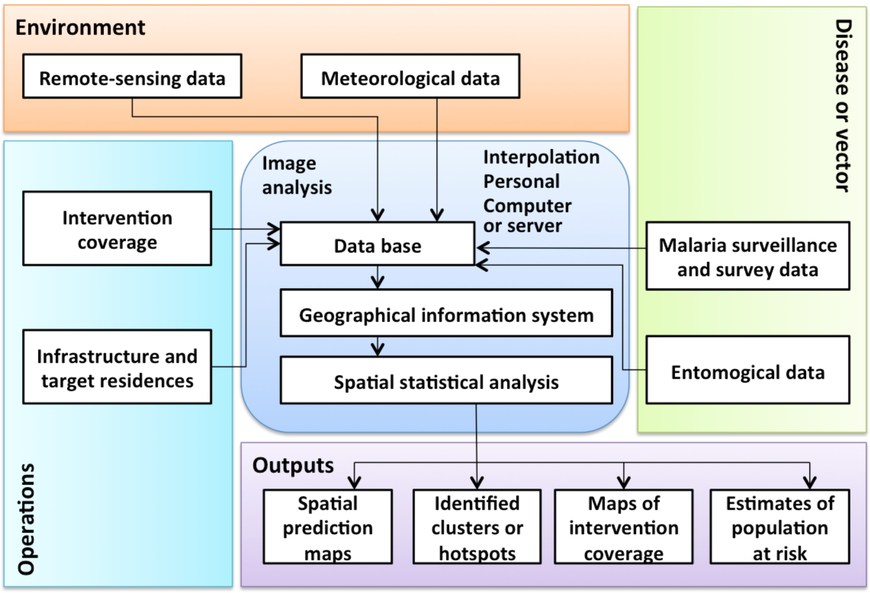



**Figure 7 Framework of geospatial science applied to disease mapping (created from [52])**

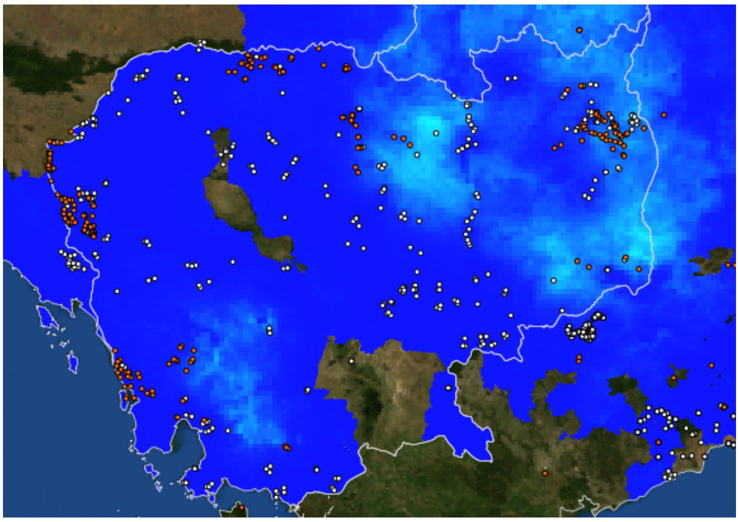


**Figure 8 Spatial distribution map of *P. falciparum* in Cambodia (created from [53])**

Remote sensing technologies are the powerful tools that can be used to identify hotspots and to investigate malaria epidemiology [7]. Several environment-related indices calculated from remote sensing data, such as normalized difference vegetation index (NDVI), normalized difference water index (NDWI) and topological wetness index (TWI) have been used to predict regional malaria endemicity [56-59]. Climate also is closely related to the risk of malaria [60-61]. Cohen et al. created fine-scale risk maps of both high endemic and low endemic seasons from routine collected individual case data combined with environmental indices calculated from remote sensing data [62]. Other variables, such as the distance from health facilities, socioeconomic status and/or status of containment interventions (for example, insecticide-treated net distribution and indoor spraying of residual insecticide) have been incorporated into risk models to predict regional malaria endemicity [62-63]. It is important to take human interactions with environment into the model when describing the risk as malaria is transmitted through the interaction with vectors lurking in the surrounding environment. In fact, several behavioral factors and human reactivity to this disease are incorporated in the mathematical model for the prediction of malaria transmission [64-66]. Ellis et al. proposed the concept of anthropogenic of the world (Figure 9) [67]. This concept is, in short, the conceptual thinking of "putting people in the map", meaning environmental factor together with anthropogenic factor such as population density may explain how the people use or interact with surrounding environment, which is expected to relate to the geographically observed phenomenon. It is also important to consider the influence of the human mobility for the malaria transmissions. Using spatially explicit mobile phone data and malaria prevalence information, Wesolowski et al. projected the source and sink of malaria parasites in Kenya. By that, they could identify the dynamics of human carriers that may drive parasite importation between regions [6].

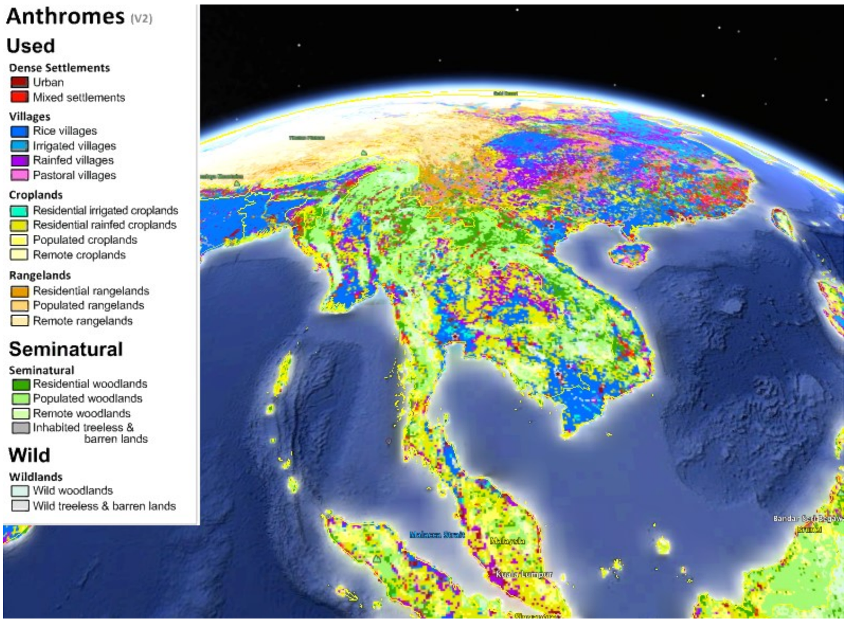



**Figure 9 Anthropogenic biome visualized in the Google Earth© (created from [67])**

## 2.3 Limitations of existing malaria mapping and measurement approach

Despite these advancements, the decrease in the prevalence of malaria causes several situational changes, which indicates that the risk mapping approach needs to be adjusted. Details of the hypothetical approaches to address these issues are discussed in the next chapter.

### 2.3.1 Measurement of the malaria risk

In terms of the measurement of malaria disease burden when malaria becomes rare, it becomes increasingly difficult to detect ongoing transmission monitoring by PR [11]. Since, these situations are the important steps toward malaria elimination, there is an important need for examining the modeling method of disease burden under low to moderate transmission settings. Annual Parasite Incidence (API) can be a reliable measure for reporting new malaria infections under these settings supported by good reporting systems [68]. However, intensive focused screening method indicated that, in low-transmission settings, not a few malaria cases are asymptomatic, which makes it difficult to identify all the cases by passive surveillance systems [41, 47].

### 2.3.2 Accessibility for fine-scale data

Under low transmission settings where few infectious cases reported, the sample size required to estimate and spatially predict infection prevalence become very large, and such information usually cannot be obtained on a fine-scale. Instead, cross-scale predictions using the data collected on a coarser scale can be performed using the Bayesian modeling framework [63].

### 2.3.3 Assessment of the effectiveness of combinations of interventional measures

At present, there are a few examples of investigations for the benefits of combining different vector control measures, but further studies are needed about the assessment of the effectiveness of using these combined approaches. This issue has become increasingly important, as "One-size fits all" approach is no longer applicable to the areas under the low-to-moderate transmission settings. Under the ever-changing local endemic conditions, all interventions need to be reviewed carefully and tailored for regional circumstances in an ongoing way to ensure that they remain fully effective. The progresses of malaria containment actions are expected to affect these conditions.

## 2.4 Studies of quality improvement and management of health related data in the health information system

The lack of quality health information is particularly alarming in situations with limited healthcare resources, which is observed in some less-industrialized countries. So far, several studies have demonstrated measures to improve the quality of health information in the health information systems. Design improvements in data-collection tools, such as the ones in the user interface of data-entry systems, were proven effective in enhancing user efficiency and reducing the data-entry errors [69-70]. In addition, several studies have addressed the data cleaning of collected data in database literature works [71]. The survey design is extensively discussed. A good form design is a cornerstone for obtaining high-quality data [72]. Various measures have been taken to not only improve the surveillance tools, but also address process and operational improvements such as introducing the double-data entry [73], field worker training, and periodical site monitoring [74]. However, some of these individual measures are sometimes considerably expensive to be incorporated uniformly in the system; as such, it is necessary to analyze the aspects of the system as a whole and identify appropriate points of intervention. The health metrics network provided frameworks and standards for an integrated approach to enhance the capability of strengthening the health information system [10]. The technical demands are guided into the continuous system management to cope with the changing environment. The mathematical programming approach to support healthcare resource allocations of the healthcare SoS was previously studied [75].

## 2.5 Limitations in current approaches for the health information systems

In spite of substantial efforts and advancements in the previous studies, the practical method for addressing the system transformation and its effect on the system management is required given that the SoS itself transforms over time. The application of graph theory for SoS engineering has been extensively discussed in recent years [76]. The knowledge and approaches generated from this attempt can help explore ideas on optimizing SoS and calculating their complexity. Therefore, this study examines an approach to model and analyze the transforming health information SoS using the process, architecture, and the risk associated with the environment.

## 2.6 Summary of literature review and requirement for the system

As explained in the previous sections in this chapter, several limitations exist in the previous studies that need to be addressed to realize the proposed system. In other words, they can be translated into the requirements of the system listed in below.

i. *The system shall sustain its operation in the context of low-to-moderate malaria transmission setting.*
ii. *The system shall be cost-effective with reasonable operational cost by taking full advantage of existing resources.*
iii. *The system shall take dynamic situation and its transformation and provide necessary information on a timely manner.*

First, the final access phase has not yet been extensively discussed in the previous study, where the individual countries seek to sustain the use of a technology for long-term prevention, control, or eradication of related disease. The aim of this study is to develop the continuous management cycle of improving healthcare resource deployment and health information system strengthening. Hence, this access phase is an important area of application. In this particular access phase, the contexts of the system need to be fully taken into account to make sustainable. They could vary and be depending on the system of interest. In the case of Cambodian malaria surveillance system, the decrease in the prevalence of malaria causes several situational changes such as the measurement of malaria and sustaining the access to fine-scale data, which indicates that the risk mapping approach needs to be adjusted. Furthermore, it is commonly observed that securing the same degree of investment for malaria containment as that obtained during the highly endemic period becomes more difficult. Hence, the system should be cost-effective with reasonable operational cost. Under the low-to-moderate malaria transmission setting, the situation may dynamically change in accordance with various factors such as the progress of containment interventions and environmental changes. The health information system is also reformed by such changes and requirement of stakeholders. Therefore, the system needs to take these situational changes and provide necessary information on a timely manner. In spite of substantial efforts and advancements in the previous studies, the practical method for addressing the system transformation and its effect on the system management has not been extensively discussed particularly in the area of system-of-systems engineering.

# 3. DISEASE RISK MAPPING CONSIDERING ENVIRONMENTAL CONTEXT DISPARITIES USING ROUTINE SURVEILLANCE DATA

In this chapter, we demonstrate the development and evaluation of the spatiotemporal malaria risk model corresponding to the environmental context disparities. Subsequently, fine-scale maps using the spatiotemporal malaria risk model were created. While risk distribution modeling and a mapping approach are effective tools to assist with the efficient allocation of limited healthcare resources, these methods need some adjustment and reexamination in accordance with changes occurring in relation to malaria elimination activities. Limited available data, fine-scale data inaccessibility (for example, household or individual case data), and the lack of reliable data due to inefficiencies within the routine surveillance system, make it difficult to create reliable risk maps for decision-makers or healthcare practitioners in the field. Furthermore, the risk of malaria may dynamically change due to various factors such as the progress of containment interventions and environmental changes. To make the continuous cycle of disease mapping in the system, this complex and dynamic nature of situations needs to be addressed.

## 3.1 Hypothetical questions and proposed approach

Based on the current situations and findings from previous studies we propose following hypothetical questions:

(1) *Can we use standardized morbidity ratio (SMR) calculated by API in the spatial risk prediction model as an appropriate measure of disease burdens?*

(2) *Can we predict the fine-scale risk better if human interactions with surrounding environment, i.e. environmental context, are considered?*

(3) *Can we present the expected outcome of interventional measures by incorporating containment status indicators into the risk prediction model?*

Here, we applied a mathematical modeling approach for SMR calculated by API using routinely aggregated surveillance reports and variables related to human interactions to surrounding environment to create spatial risk distribution maps in fine-scale of two provinces (Pailin and Preah Vihear) in western Cambodia where the artemisinin resistance was previously reported. In addition, we incorporated the combinations of containment status indicators into the model, by which the regional heterogeneities of the relationship between containment status and risk can be visually represented for the efficient healthcare resource allocations and intervention planning considering temporal descriptions of regional malaria endemicity. To address the complex and dynamic nature of situations in low-to-moderate malaria transmission settings, the model was expanded to consider spatial and temporal changes. We build a spatiotemporal model of SMR of malaria incidence by employing the hierarchical Bayesian frame to fit the transitioning malaria risk data onto the map. The model was set to estimate the SMRs of every study location at specific time intervals within its uncertainty range. Based on the spatiotemporal risk model developed here, we also estimated and visually presented the priorities of constituent bodies involved in the routine surveillance network, that is, the relative weights of network priorities for relevant constituents, using graph theory analysis. The aim of this analysis is to help understand the transitional complexities existing in the system, in support of better informed decision-making for more efficient resource allocation and intervention planning, through the consideration of spatiotemporal description of regional endemicity.

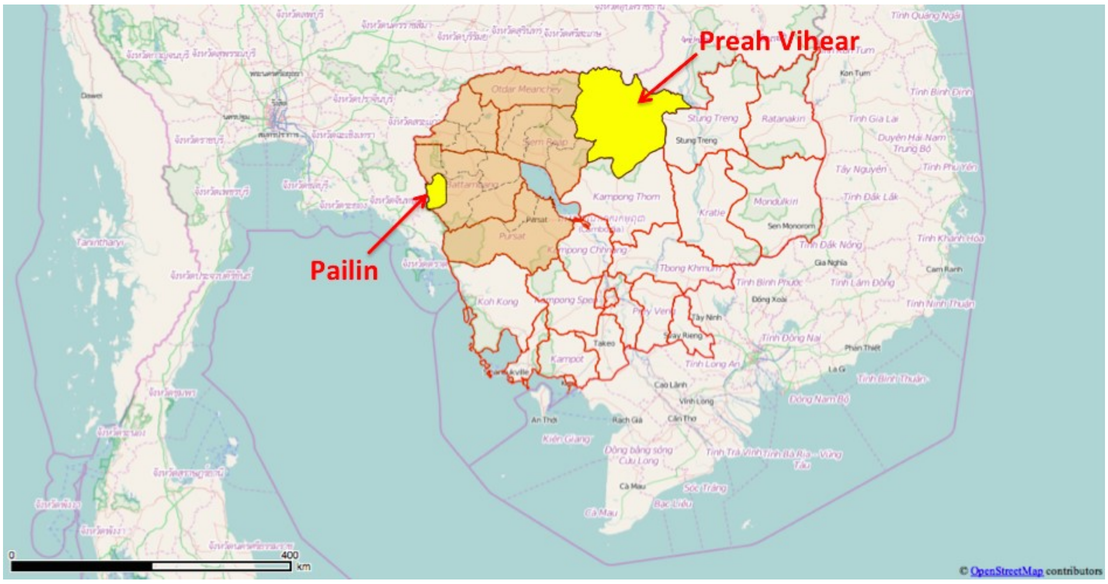



**Figure 10 Map of the research area**

Open Street Map© was used to create this map

## 3.2 The environmental context disparities

Human interactions with surrounding environment are the important factors affecting the risk especially for communicable diseases transmitted through infectious vectors. Figure 11 shows an applied case of the epidemiologic triad of disease causation [77] for the malaria case. This triad consists of an external agent (Malaria), a host (Human) and an environment in which host and agent are brought together, causing disease. Vector plays as the transmission carrier of malaria parasite that does not present malaria symptoms.

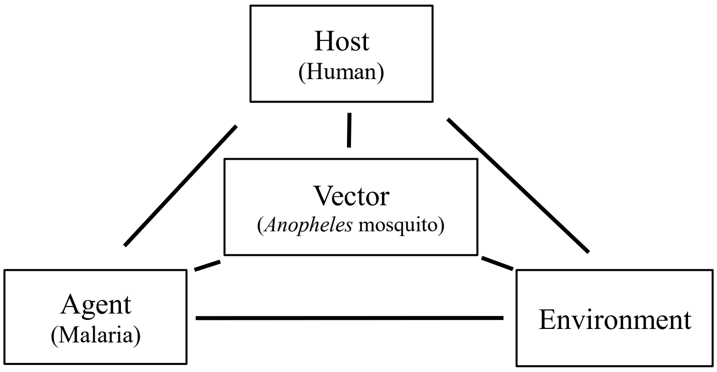



**Figure 11 An applied case of the epidemiologic triad of disease causation of malaria**

Based on this concept, the environmental context surrounding human communities, i.e. how the environmental features exist and interact with people, is an important factor affecting the risk of this infectious disease. Thus, disparities in environmental context among communities could explain the extent of disease burden across areas. In this study, we defined this concept as "**The environmental context disparities**" as a key factor for the development of malaria risk model.

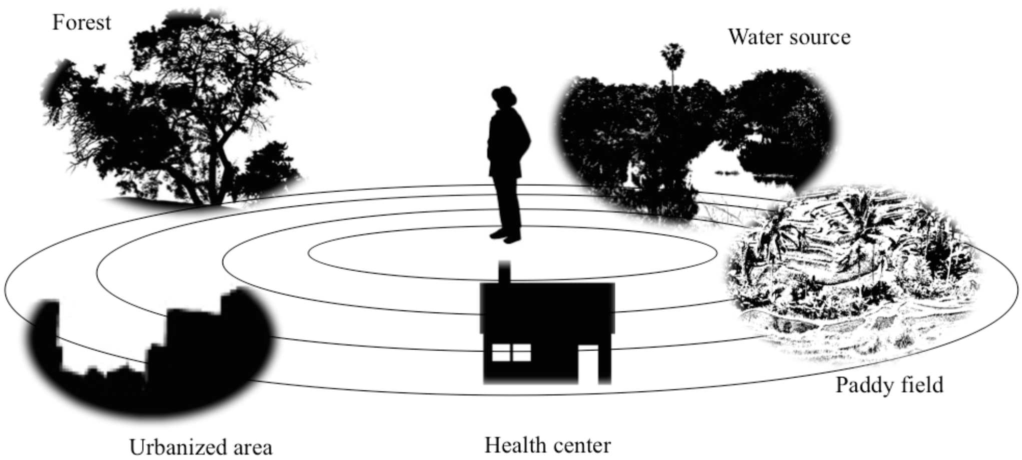



**Figure 12 Concept figure of environmental context**

## 3.3 Materials and methods

### 3.3.1 Malaria data collection

Malaria data were collected from Cambodia malaria bulletin report from 2010 to 2013 [78-79]. This dataset was build from case reports collected through the effort of the Malaria Information System (MS) and the national facility-based Health Information System (HIS) using a common coding system [80]. It contains API (per 1000 people) in each health operational district (HOD) for two malaria species, *P.falciparum* and *P.vivax*, reported by healthcare facilities or village malaria workers. The SMR, standardized mortality or morbidity ratio, is expressed as a ratio or percentage of quantifications compared with the general population of interest (equations 1, 2) [81].

$$SMR = \hat{\theta}_i = \frac{o_i}{e_i} \qquad (1)$$

$$e_i = \sum_k n_{ik} P_k \qquad (2)$$

where, $o_i$ is the observed number of cases in *i* area, $e_i$ is the expected number of cases in *i* area, $n_{ik}$ is the population in *k* age group in *i* area, and $P_k$ is the incidence of clinical cases in *k* age group in the reference population. $e_i$ was estimated by multiplying the population and reported incidence and aggregating them for each age group in 10 provinces in western Cambodian [82]. Since, the API was reported incidence per 1,000 people, SMR, $\hat{\theta}_i$ in *i* district, was calculated by dividing API by $e_i$ per 1,000 people. Assuming small observed case numbers and relatively large dispersions under the low-to-moderate transmission settings, the observed case count data $o_i$ can be assumed to follow the negative binomial distribution, $o_i \mid \mu_i$, where $\mu_i$ is the corresponding distribution mean and $\rho$ is the scale parameter (equation 3). Then, by transforming equation 1, $\mu_i$ can be derived by multiplying $e_i$ and the relative malaria risk, $\hat{\theta}_i$ (equation 4) [83]. Hence, SMR can be used for estimating the case number of target area, which is also useful for informed decision-making of healthcare resource allocation.

$$o_i \mid \mu_i \sim \text{NegBin}\,(\mu_i, \rho) \qquad (3)$$

$$\mu_i = e_i \hat{\theta}_i \qquad (4)$$

Considering the small numbers of observed cases compared with population size under the low-to-moderate malaria transmission setting and thereby raising a concern for the modifiable areal unit problem in geographical analysis [84-85], the SMR for each HOD was smoothed using the empirical Bayesian method (EBSMR) [86] to adjust the influence of different population size in area units. EBSMR was calculated by equation 5-7, given that $\hat{\theta}_i$ follows gamma distribution (equation 8) and observed $o_i$ under $\theta_i$ follows Poisson distribution (equation 9).

$$EBSMR = \hat{\theta}_i = e[\theta_i | o_i, e_i] = \frac{o_i + \upsilon}{e_i + \alpha} \qquad (5)$$

$$\frac{\hat{\upsilon}}{\hat{\alpha}} = \frac{1}{n}\sum_{i=1}^{n} \hat{\theta}_i \qquad (6)$$

$$\frac{\hat{\upsilon}}{\hat{\alpha}^2} = \frac{1}{n+1}\sum_{i=1}^{n} \left(1 + \frac{\hat{\alpha}}{e_i}\right)\left(\hat{\theta}_i - \frac{\hat{\upsilon}}{\hat{\alpha}}\right)^2 \qquad (7)$$

$$\theta_i \sim Ga(\upsilon, \alpha) \qquad (8)$$

$$o_i | \theta_i \sim Po(\theta_i, e_i) \qquad (9)$$

### 3.3.2 Environmental and non-environmental anthropogenic covariates

The covariates that were incorporated into the modeling framework are described in Table 2. The NDVI, the NDWI, and the land surface water index (LSWI) were calculated from Terra-MODIS 8-day composite data (http://LPDAAC.usgs.gov) from 2010 to 2013. Because EBSMR was represented as yearly average, these environmental variables were averaged to the mean values for each year. NDVI, an index correlating with the extent of vegetation and used for forest monitoring was calculated using the reflectivity of red in visible range $(R)$ and near infrared radiation range $(IR)$ collected by satellite sensor (equation 10). For the MODIS satellite, $IR$ corresponds to band 2 and $R$ corresponds to band 1 (equation 11).

$$NDVI = \frac{IR - R}{IR + R} \quad (10)$$

$$NDVI = \frac{(Band2) - (Band1)}{(Band2) + (Band1)} \quad (11)$$

In the same manner, NDWI and LSWI, indices of water, can be calculated by combinations of the reflectivity of different wave lengths (equations 12, 13).

$$NDWI = \frac{(Band2) - (Band5)}{(Band2) + (Band5)} \quad (12)$$

$$LSWI = \frac{(Band2) - (Band6)}{(Band2) + (Band6)} \quad (13)$$

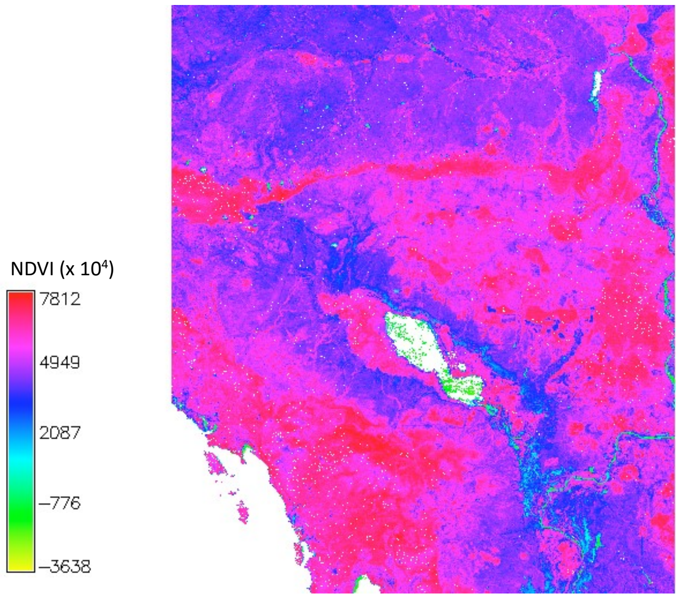



**Figure 13 Example of NDVI calculation using MODIS satellite data**

Note: NDVI values were multiplied by $10^4$.

The digital elevation model at 30-m resolution was extracted from ASTER GDEM database (http://gdem.ersdac.jspacesystems.or.jp) [87] and used to estimate the altitude. The TWI was calculated using this altitude model and estimated by the method described in previously [88]. Considering interactions between surrounding environment and people in the malaria transmission process, we extracted data from multiple surrounding circular buffers with different radius distances (i.e., for each 1km from 1-5 km) from villages. Distance and surrounding circular buffers were generated by the Quantum GIS software. Environmental covariates extracted from each village were aggregated to the district level to reflect the overall condition of target district. As the number of villages directly relates to the aggregated value, they were taken averages by the number of villages in each HODs. These data, which could potentially indicate the human interactions with the surrounding environment, were compared by calculating the correlation and coefficient of determination for the models. Because temperature can influence the ecology of mosquito breeding habitats, and therefore malaria transmission [60], we collected the *Plasmodium* temperature suitability index (*Pf*TSI) [89] from Malaria Atlas Project database [90]. Rapid urbanization is related to changes in the risk patterns of malaria transmission compared with rural sparsely populated areas [91-92] and the susceptibility of these two different populations can be influenced by the types of implementations of containment actions that are implemented. Population density per $km^2$ was calculated as a variable reflecting the extent of urbanization, using records in the Cambodian Malaria Bulleting divided into the areas of each HOD. Furthermore, we used the reported proportion of sufficient ownership of long lasting insecticide-treated nets (LLIN) [82] and treatment failure rate of artemisinin ($TF_{rate}$) [38] as containment status indicators. $LLIN_{suf}$ is defined as the proportion of households in which distributed mosquito net covers no more than two persons per net. Because no geographical localities could be obtained for these indicators, they were aggregated to the provincial level and incorporated in the model development.

**Table 2 Variables used to build the modeling framework to estimate EBSMR**

| Category | Variable | Data source | Data collection |
|---|---|---|---|
| Vegetation | NDVI | Terra-MODIS 8-day composite data 2010-2013 | Extracted mean value from 1, 2, 3, 4, 5 km surrounding circular buffer from each populated village |
| Water | NDWI | Ditto | Ditto |
| | LSWI | Ditto | Ditto |
| Geography | TWI | Digital elevation model at 30 m resolution from ASTER GDEM database [87] | Ditto |
| Temperature | *P. falciparum* Temperature suitability index (*Pf*TSI) | Malaria Atlas Project database [90] | Averaged to mean value for each HOD |
| Population | Population density (/$km^2$) | Cambodia malaria bulletin report 2010-2013 [78-79] | Population record divided by total areas of each HOD |
| Vector control | Sufficient ownership of LLIN [a] | Cambodia malaria survey 2010 [82] | Used the values reported at each provincial level |
| Treatment | Treatment failure rate by artemisinin combination Therapy [b] | National Center for Parasitology, Entomology and Malaria Control [38] | Ditto |

[a] Proportion of household in which distributed mosquito net covers 2 persons or less per net.

[b] Test positive for *P. falciparum* on day 28 or day 42

EBSMR, Standardized morbidity ratio estimated by empirical Bayesian method; NDVI, Normalized difference vegetation index; NDWI, Normalized difference water index; LSWI, Land surface water index; LLIN, Long lasting insecticide-treated net; Topographical wetness index; HOD, Health operational district

### 3.3.3 Spatial risk distribution modeling

The relationship between EBSMR ($\hat{\theta}$) and spatial covariates was modeled using a generalized linear regression model as a function of the *N* predictive variables $(X, Z)$, given that the logarithmic $\hat{\theta}$ follows the Gaussian distribution.

$$\hat{\theta} = e^{\lambda} \tag{14}$$

$$\lambda = \alpha + \sum_{N} \beta_N X_N + \sum_{N} \gamma_N Z_N + \varepsilon \tag{15}$$

where $\alpha$ is the model intercept, $\beta$ is the parameter associated with environmental covariates $X$ and $\gamma$ with non-environmental anthropogenic covariates $Z$. The maximum likelihood of observed data provided to the model and the input predictors were calculated based on this modeling frame (equation 14 and 15). Data modeling was conducted on the district level scale. For the model fitting, either maximum likelihood or Markov Chain Monte Carlo (MCMC) methods can be used. We first used the maximum likelihood method to examine the predictor variables and then, based on the results, we used the MCMC method in the Bayesian modeling frame to estimate the uncertainty about the relationships represented by $\alpha$ and $\beta$, and $\gamma$ (equation 15) and cross-scale predictions. The models were fitted using the R software (https://www.r-project.org). Predictor variables were entered into the initial models in a stepwise manner to identify the variables to be incorporated into the model. This approach was repeated until all remaining variables in the final model were significant at $\alpha$=0.05. An MCMC sampler in the JAGS framework [93] was used for the Bayesian model fitting. Three MCMC chains with 50,000 iterations as burn-in and 30,000 iterations thinned every 30 were stored as parameter estimates. Convergence of the model was examined by Gelman-Rubin diagnostics [94] and visual assessment of trace plots of chains (see appendix).

### 3.3.4 Mapping and validation

The fitted model was applied in conjunction with spatial covariates extracted from the location of each village to estimate the village level SMR. This process can be considered as the disaggregation process of aggregated environmental covariates once used for modeling at the district level scale. Values of estimated village level SMR were used as skeletons of the spatial interpolation. Realized values calculated by spatial interpolation methods were plotted in each 250 x 250 m spatial grid. We created maps that visualized the risks of two provinces in western Cambodia, Pailin and Preah Vihear, in western Cambodia by the inverse distance weighed method (IDW) and ordinal kriging interpolation of estimated SMR at each village. To evaluate the accuracy of the cross-scale prediction from the model, the predicted SMR was compared with geocoded case data for Pailin [95] and Preah Vihear [96] collected from Malaria Atlas Project database [90] using Spearman's rank correlation [97] and Welch's t-tests for unequal variances [98]. The source data of maps created here in this study were mostly from the report from VMW and HIS and were based on the rapid diagnostic kit (RDT) and Microscopy detection. These data were selected because of the detection method (RDT / Microscopy) used and were closest to the reported period from the study period. To exclude the incidental nature for spearman's correlation with this sample data, we resampled the dataset 2,000 times with replacement to create confidence interval with the non-parametric bias corrected and accelerated percentile method [99] to assess the distribution of correlation values. Since, we aimed to provide useful information to the practitioners, visual representations of risk distributions in the maps were also validated for their agreement with those in existing risk maps and utilities of the maps for deciding target areas through interviews with healthcare providers in the regional health center and with professionals of GIS.

## 3.4 Results of the spatial risk distribution modeling

Of the 329,830 cases reported in 2011–2013, 124,888 cases reported from 18 HODs in western-Cambodia provinces (Banteay Meanchey, Battambang, Oddar Meanchey, Pailin, Preah Vihear, Pursat, and Siam Reap, in alphabetical order) were in the analysis. The SMRs in each health operational district were smoothed using an empirical Bayesian method. In contrast to the decreasing tendency the case incidence in each district, estimated EBSMRs suggested remaining or even the increasing tendencies of API in the endemic areas (Figure 14). The observed case numbers and the estimated EBSMR of each HOD through the study period are shown in figure 15 and 16, respectively. Within 5km of villages, the absolute correlation values between environmental variables (NDVI, LSWI, and TWI) extracted from surrounding circular buffers (from 1 – 5 km) and EBSMR were highest at 5 km and at 1 km for NDWI (Figure 17). Correspondingly, the Pearson correlation coefficient $R^2$ of the model differed at each distance. Thus, the data collection ranges chosen for the model was 5 km for NDVI, LSWI, TWI and 1 km for NDWI. After selecting of the spatial covariates, the final model was used to estimate the SMR of each area (adjusted $R^2 = 0.774$, Akaike information criterion AIC = 149.423). This model included NDVI, NDWI, TWI, *P. falciparum* temperature suitability index, $LLIN_{suf}$ and $TF_{rate}$. The parameter estimates for each variable are shown in table 4.

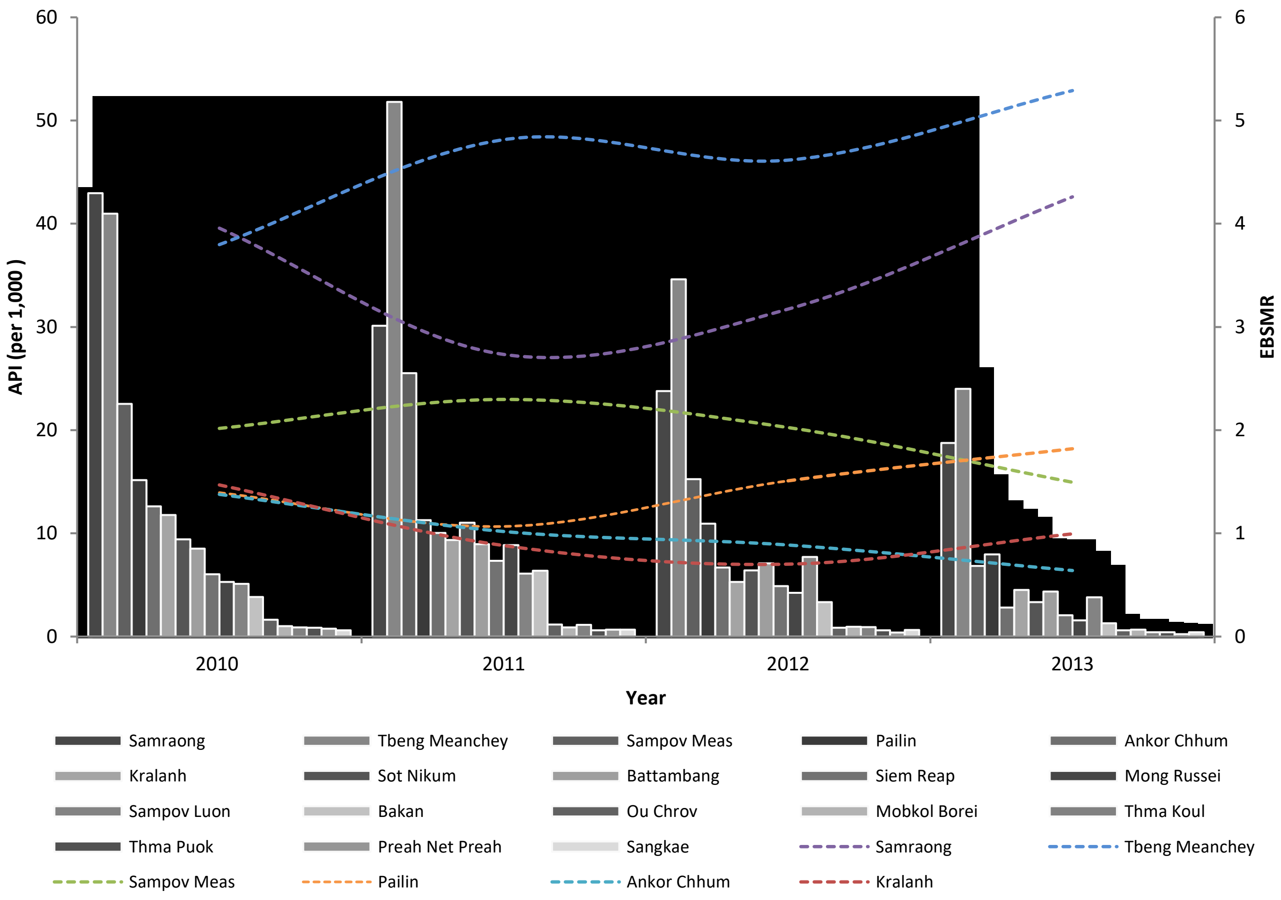



**Figure 14 Annual parasite incidence (API) of western-Cambodian health districts and empirical Bayes estimated standardized morbidity ratio (EBSMR) for 6 operational health districts with high EBSMR [a].**

Bar graph represents API in each health operational district
and dotted line represents EBSMR of 6 provinces with high EBSMR.
[a] District at higher EBSMR than 1.0

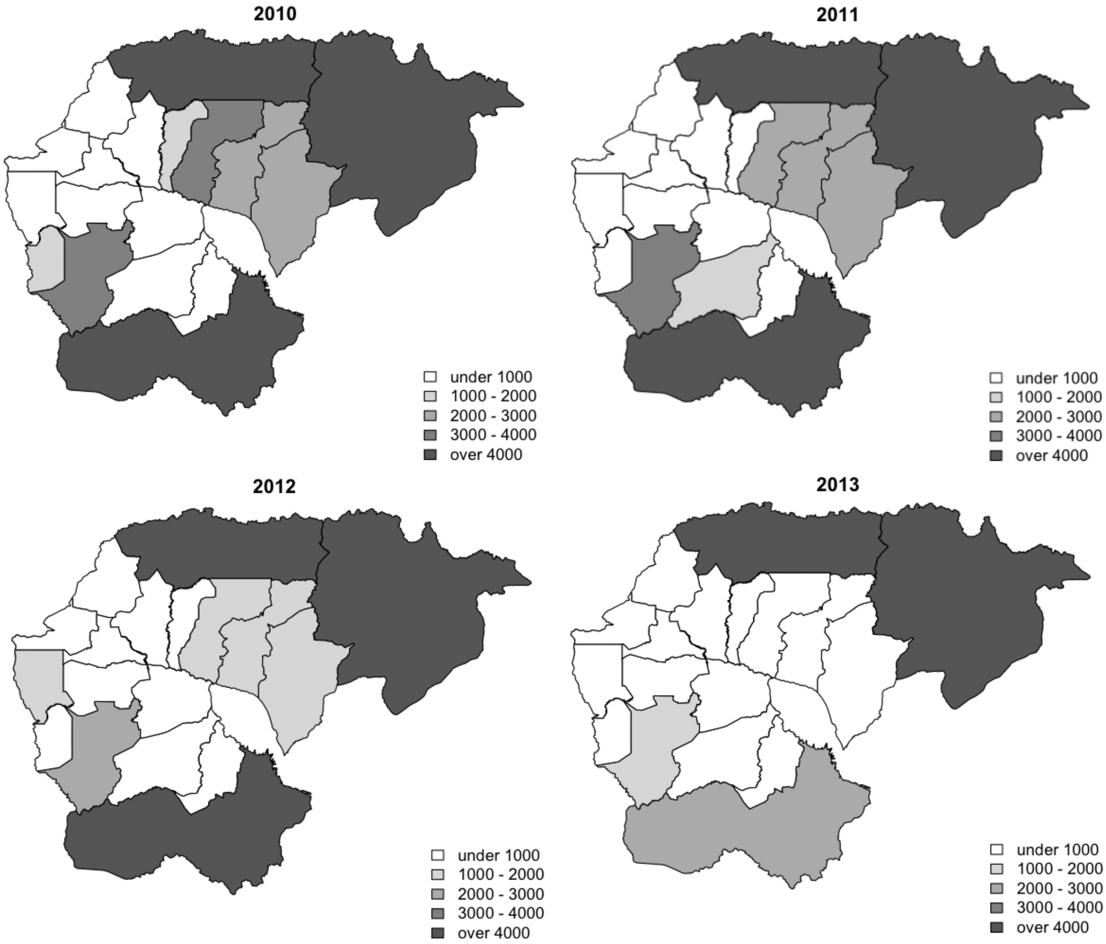



**Figure 15 Maps of annual observed case numbers of health operational districts during the study period (2010 – 2013)**

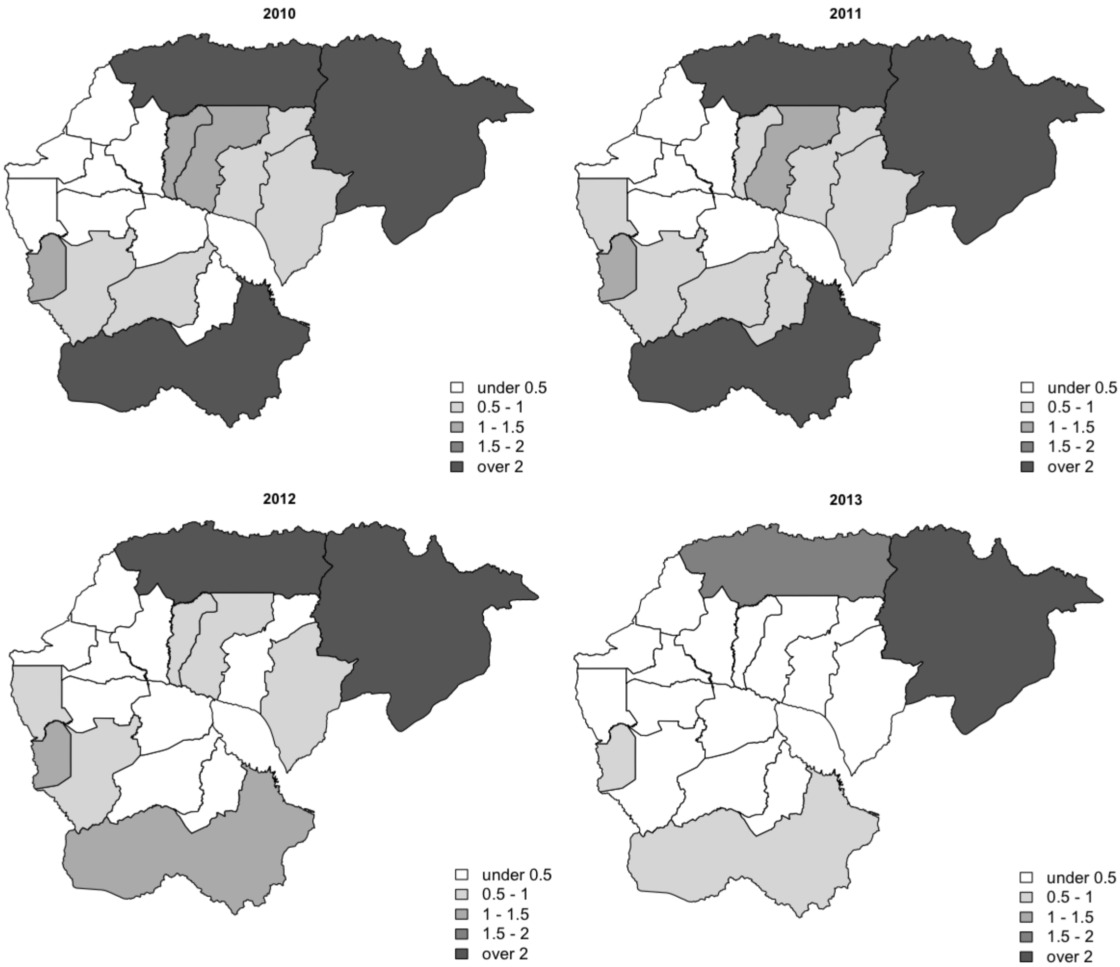



**Figure 16 Maps of annual EBSMR of health operational district during the study period (2010 – 2013).**

EBSMR, standardized morbidity ratio estimated using the empirical Bayesian method.

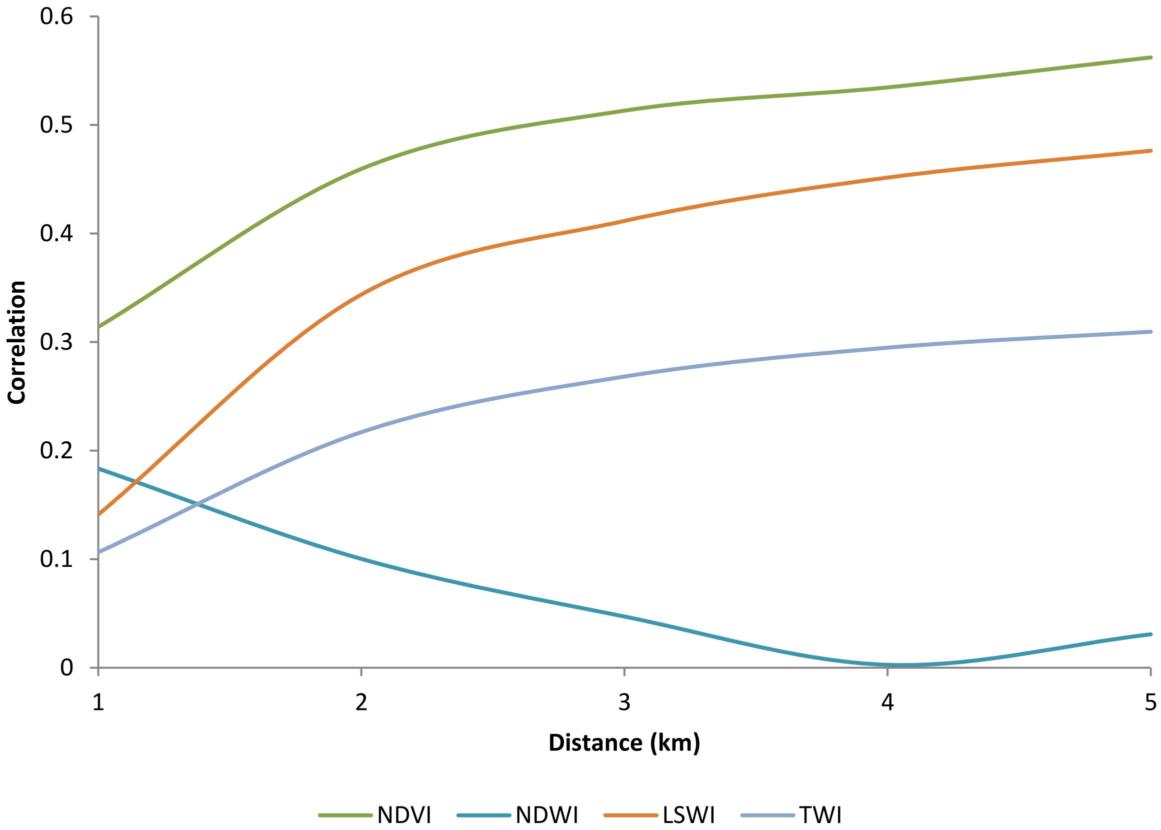



**Figure 17 Absolute correlation values between environment-related covariates extracted from surrounding circular buffer from circular buffer from populated villages and EBSMR**

Values were extracted from each 1 km distance circular buffer (1, 2, 3, 4, 5 km) from populated villages and then averaged to mean values.
EBSMR, Standardized morbidity ratio estimated by empirical Bayes method; NDVI, Normalized difference vegetation Index; NDWI, Normalized difference water index; LSWI, Land surface difference index; TWI, Topographical wetness index

**Table 3 Parameter estimates selected for the final generalized linear regression model**

| Category | Variable | Parameter estimate | Standard error | P-value |
|---|---|---|---|---|
| Vegetation | NDVI (5 km) | 7.446 | 1.947 | <0.001 |
| Water | NDWI (1 km) | -24.330 | 5.009 | <0.001 |
| Geography | TWI (5km) | -1.707 | 0.6346 | 0.009 |
| Temperature | *P. falciparum* Temperature suitability index (*Pf*TSI) | 0.0002681 | 0.0000403 | <0.001 |
| Vector control | Sufficient ownership of LLIN [a] | -0.06387 | 0.007157 | <0.001 |
| Treatment | Treatment failure rate by artemisinin combination Therapy [b] | 0.03611 | 0.008309 | <0.001 |

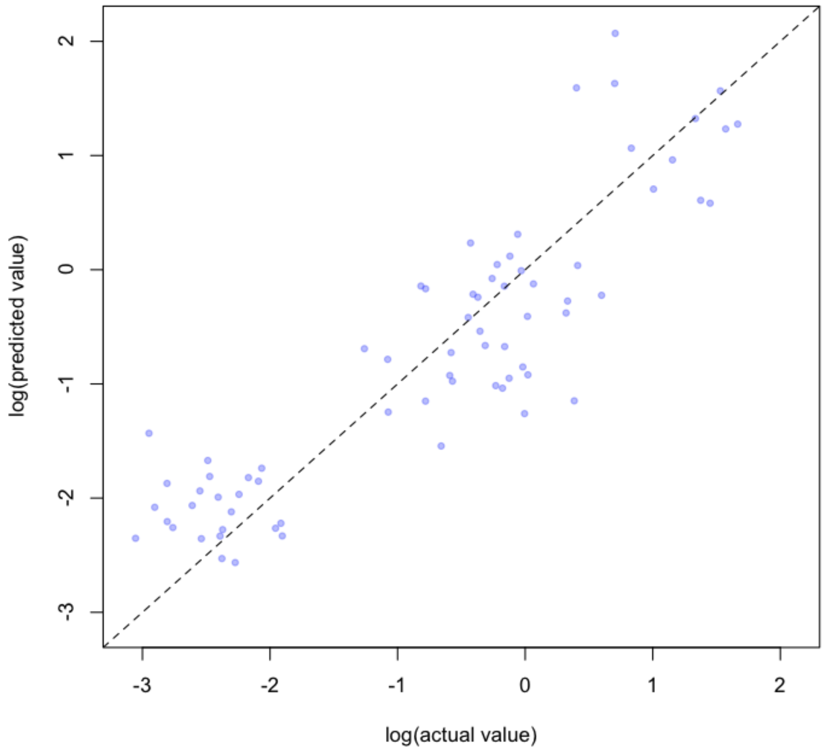
[a] Proportion of household in which distributed mosquito net covers 2 persons or less per net.

[b] Test positive for *P. falciparum* on day 28 or day 42

NDWI, Normalized difference water index; NDVI, Normalized difference vegetation index; TWI, Topographical wetness index; LLIN, Long lasting insecticide-treated net

The calibration plot of the final model indicated good fitting of the predicted and actual values (Figure 18A) and the mean absolute error (MAE) of this final model was 0.499. Figure 18B shows that 55.56% of predicted values were within the range of $\pm 0.2$, 75% were in $\pm 0.5$ and 87.5% were in $\pm 1$. Based on the information from generalized linear regression modeling, the Bayesian modeling frame was applied to estimate the uncertainty about the relationships represented by $\alpha$ and $\beta$ (equation 15). The trace plots of the Bayesian modeling frame were monitored to examine the convergence of cross scale prediction (See appendix). The model was settled with given condition for MCMC and provided the range of posterior distribution of parameters (Figure 19).

**(a)**




**(b)**

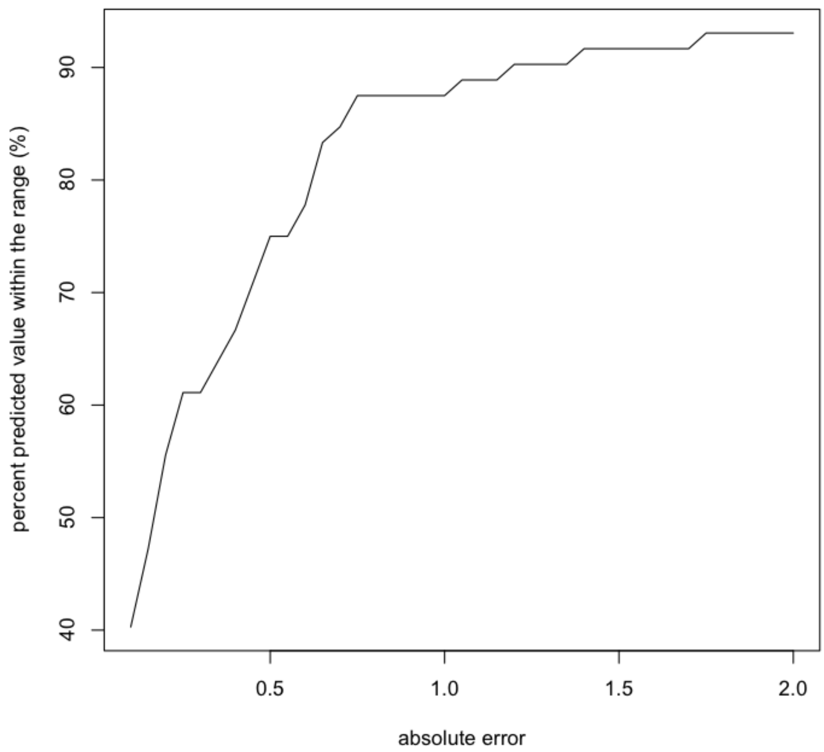
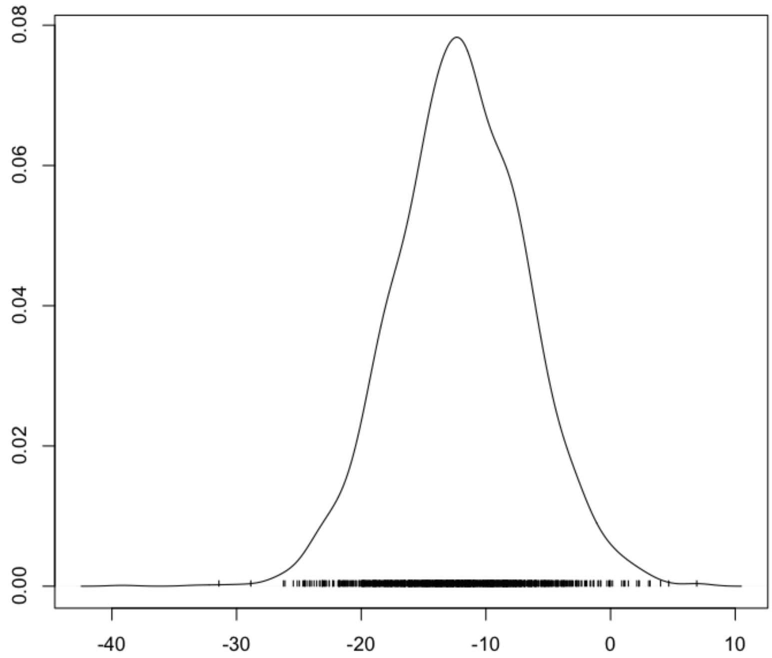


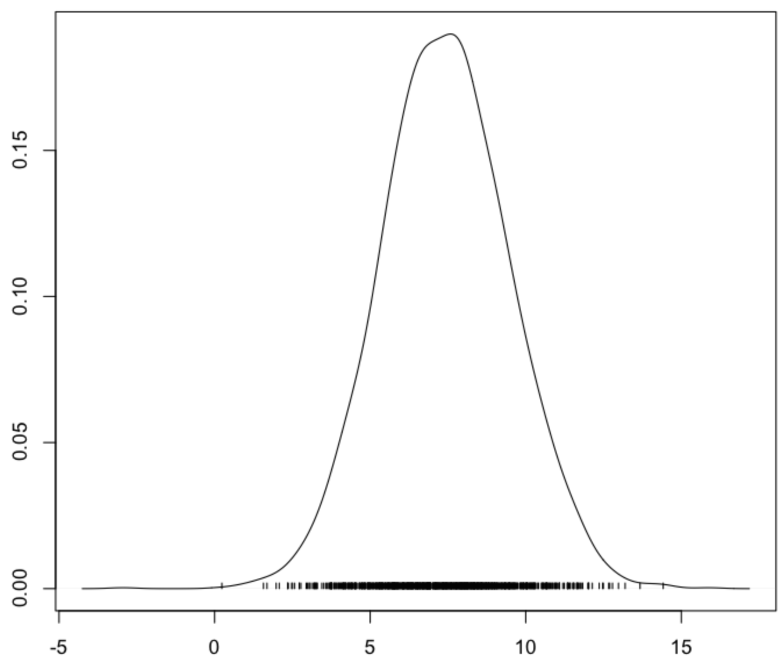

**Figure 18 The calibration plot (a) and proportion of predicted values within the range of absolute error (b) of the final model**

The dashed line in figure (a) represents 1:1 relationship of actual and predicted value.

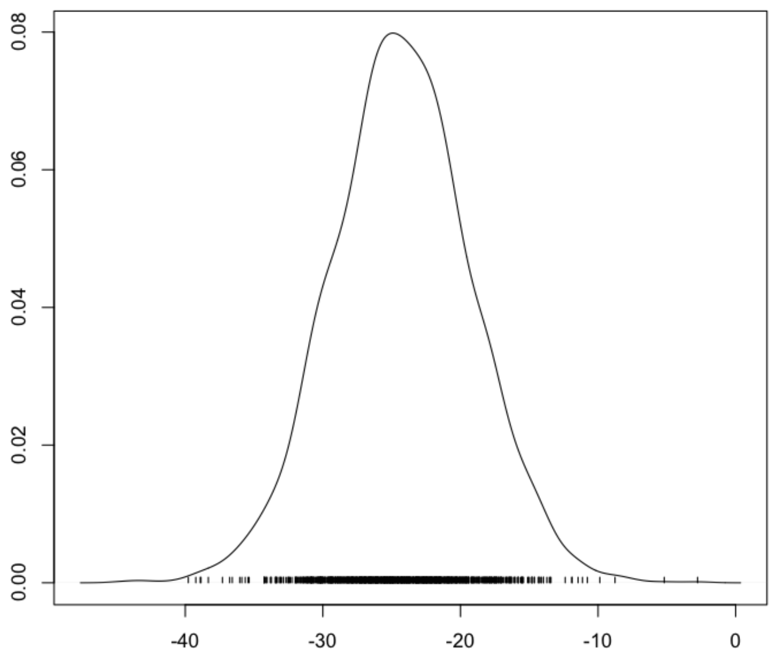
**(a)**



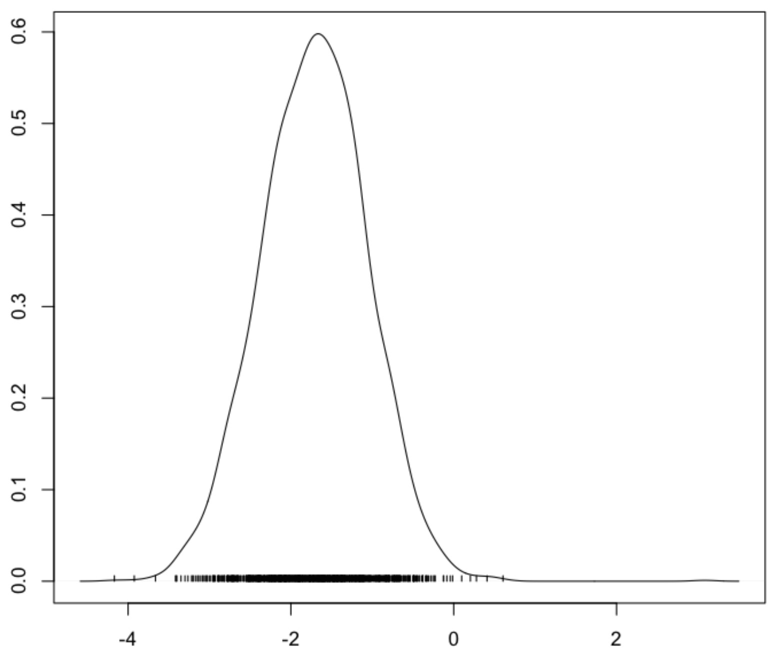

**(b)**



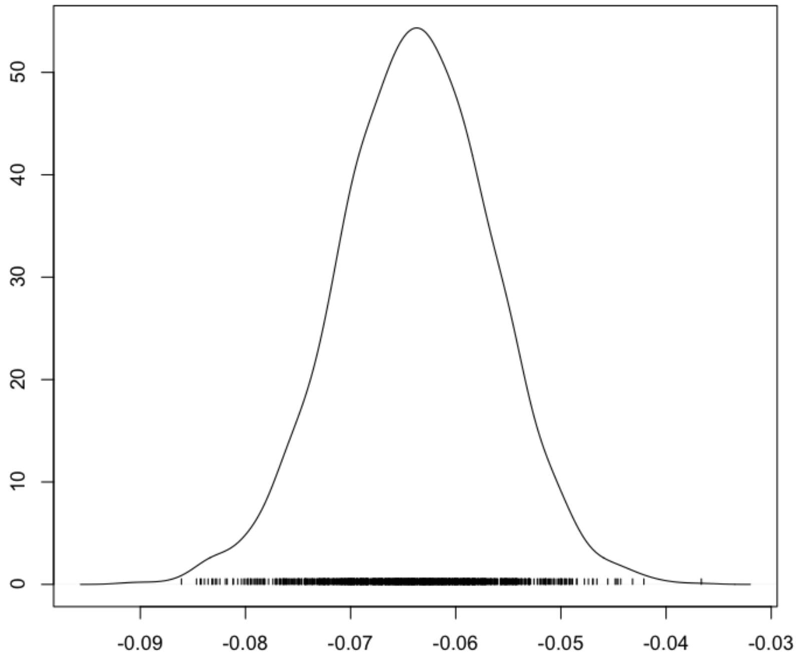
(c)




(d)

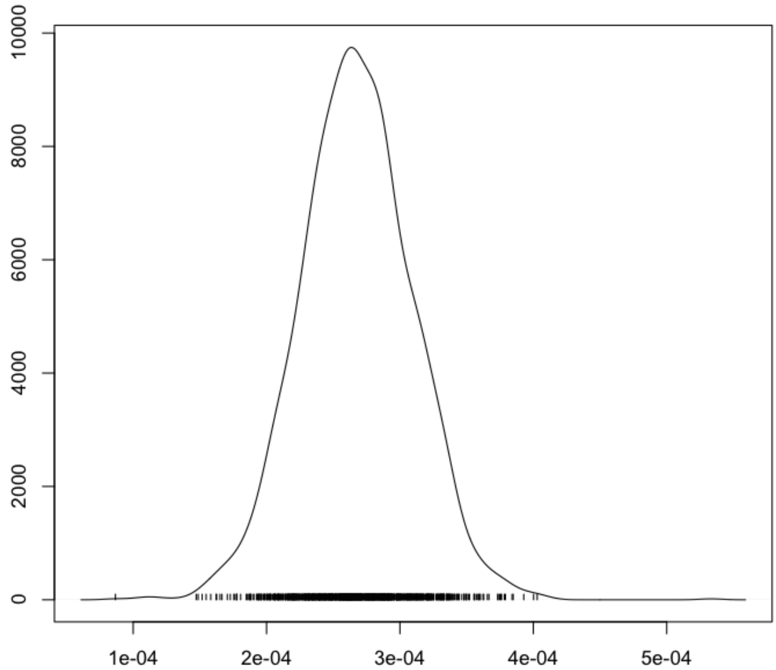


**(e)**



50
40
30
20
10
0
-0.09
-0.08
-0.07
-0.06
-0.05
-0.04
-0.03


**(f)**



10000
8000
6000
4000
2000
0
1e-04
2e-04
3e-04
4e-04
5e-04

**(g)**

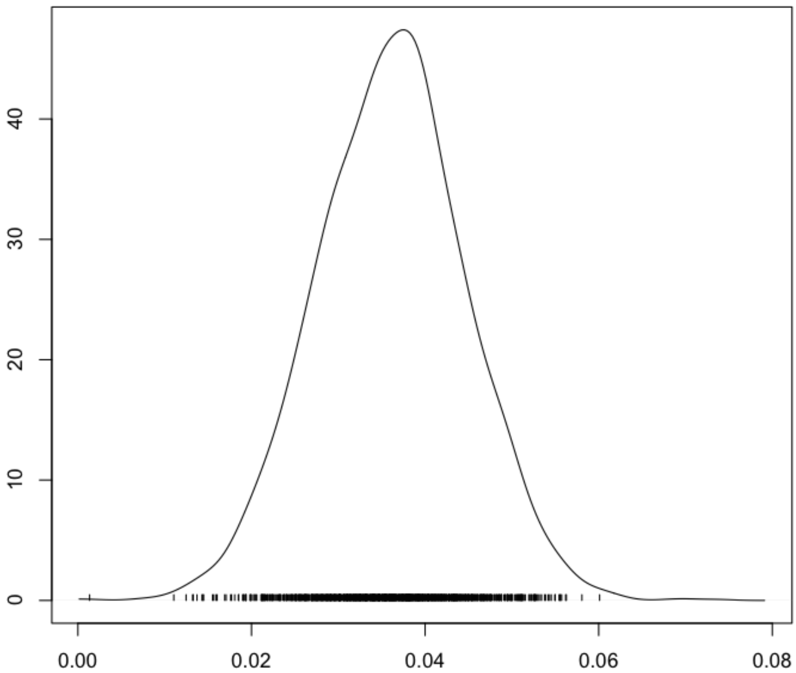


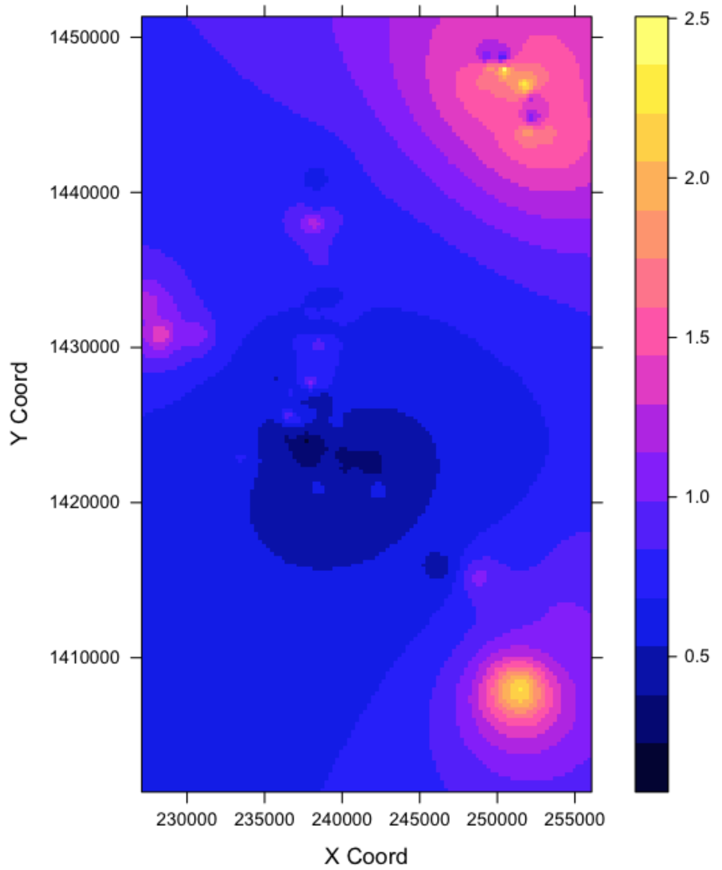

**Figure 19 Density plot of posterior distributions of each parameter**

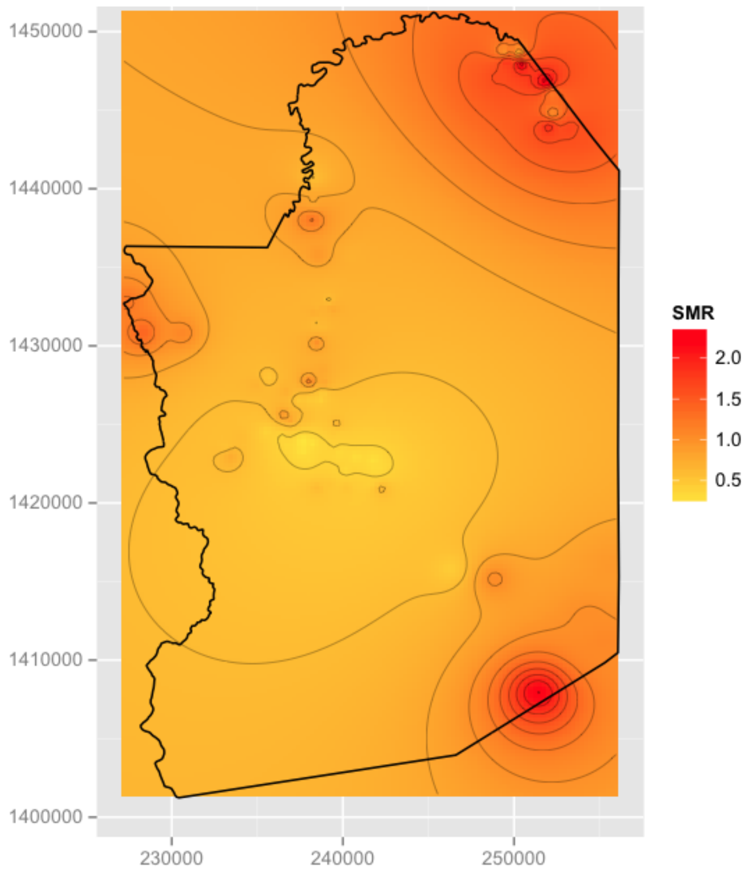
Parameter distributions for (a): Intercept, (b): NDVI, (c): NDWI,
(d): TWI, (e): LLIN, (f): Temperature and (g): TF
Horizontal axis of each graph indicates estimated kernel density of each parameter
NDWI, Normalized difference water index; NDVI, Normalized difference vegetation index;
TWI, Topographical wetness index; LLIN, Long lasting insecticide-treated net

## 3.5 Results of the fine-scale mapping and evaluation

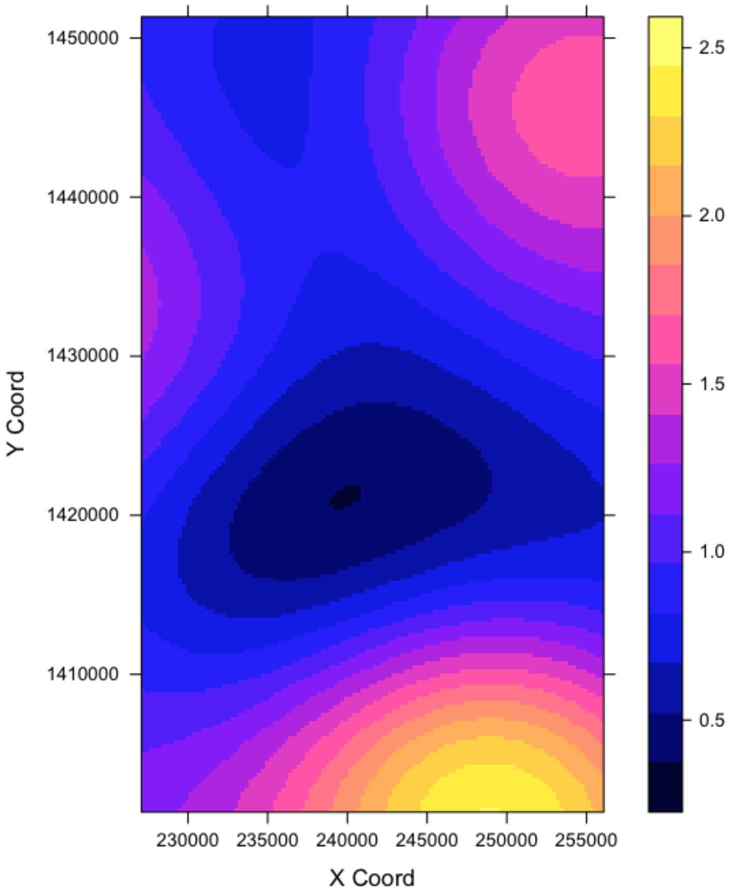
The estimated SMR for each village was calculated using the Bayesian modeling framework. Subsequently, fine-scale maps were created by the IDW method and ordinal kriging interpolation. The maps created from the predictive models for Pailin and Preah Vihear provinces are shown in figure 20. Each map represents different risk representation patterns in accordance with interpolation method used. The maps created by IDW method showed more spotted risk, which help identifying localized risky hotspots, whereas the map interpolated by ordinal kriging showed broader patterns, which provide a perspective of overall trends for optimizing healthcare resource distributions. Compared with geocoded case data, corresponding predicted values in this map showed conformity (Spearman's rank correlation; r = 0.662 with IDW and 0.645 with ordinal kriging, Welch's t-test; N.S.), which showed that the cross-scale predictions corresponded well with the actual case reports (Figure 21A). The 95% confidence intervals for the IDW and ordinal kriging methods were 0.414 – 0.827 and 0.368 – 0.813, respectively, showing a steep peak in the kernel density plot at around 0.65 – 0.7 (Figure 21B). The visual representations of hotspot in the fine-scale map created here confirmed that they were aligned with actual areas at high risk, which were identified by other sources [41, 90, 96], through the visual assessment by a number of healthcare providers and experts in the GIS. Thus, using this model, expected outcomes under given conditions of $LLIN_{suf}$ and $TF_{rate}$ were conducted. The visual representations demonstrated different patterns of expected outcomes from the combination of these two containment status indicators in respective areas (Figure 22). These simulation results could be mapped to examine the geographical effect expected from targeted containment status. Figure 23 shows an example of geographical analysis. The geographic view of the effect from each or combined containment status could be obtained from this analysis.

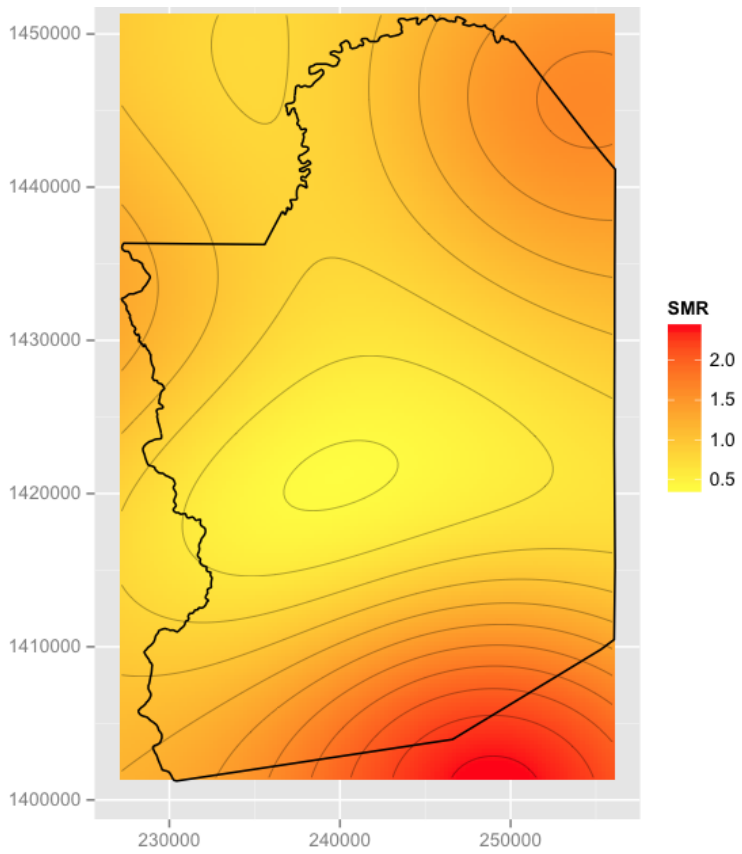
**(a) Pailin, 2010 (IDW)**



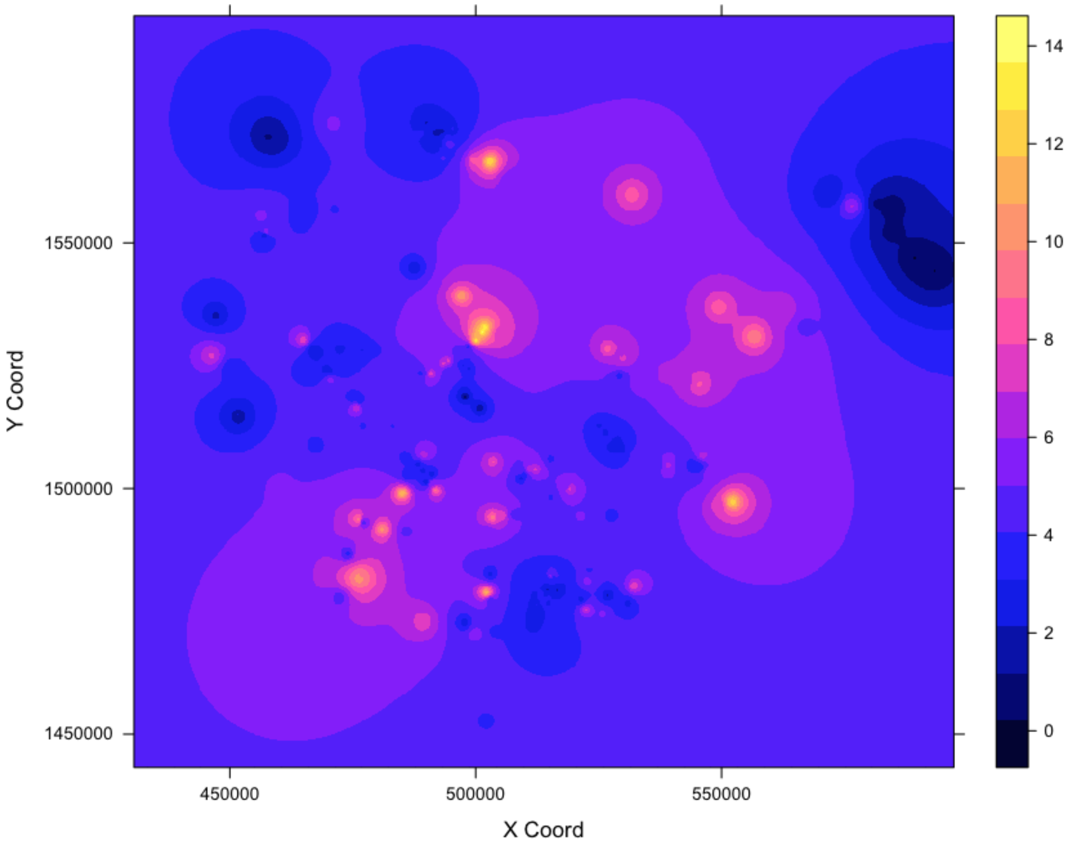



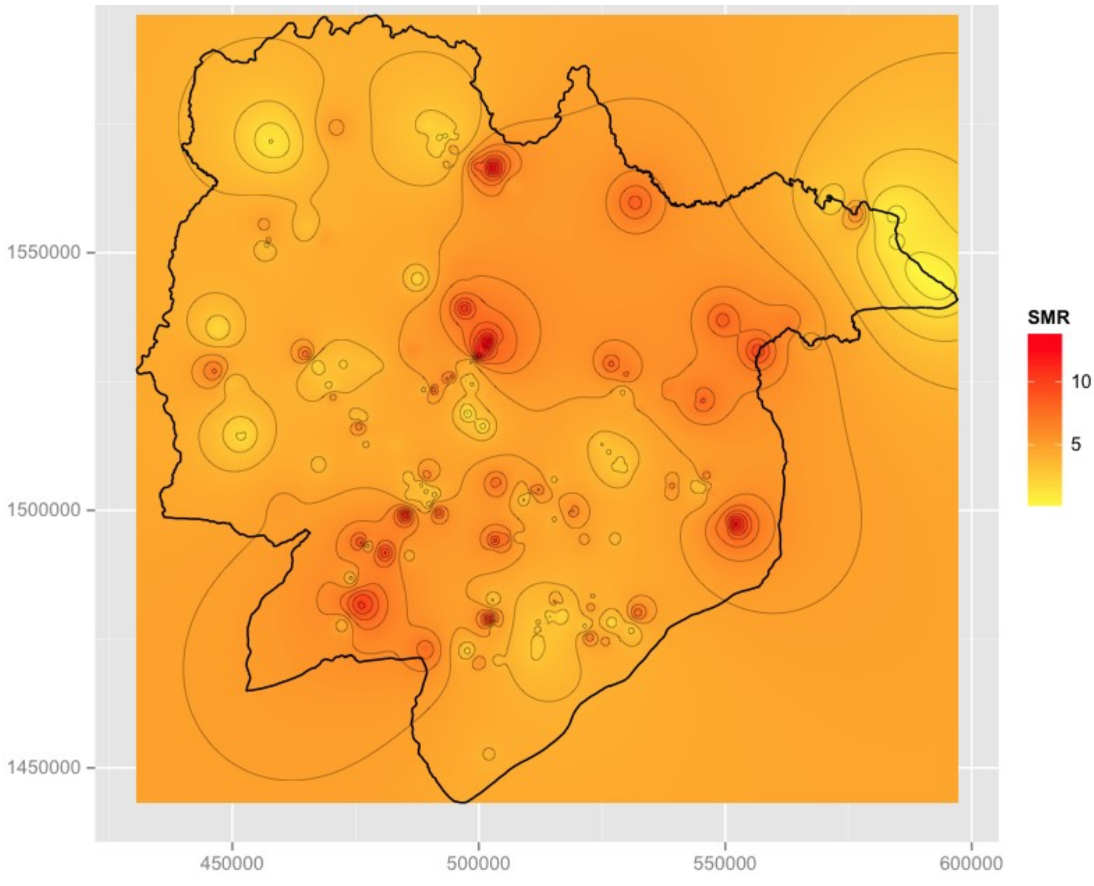
**(b) Pailin, 2010 (Ordinal kriging)**



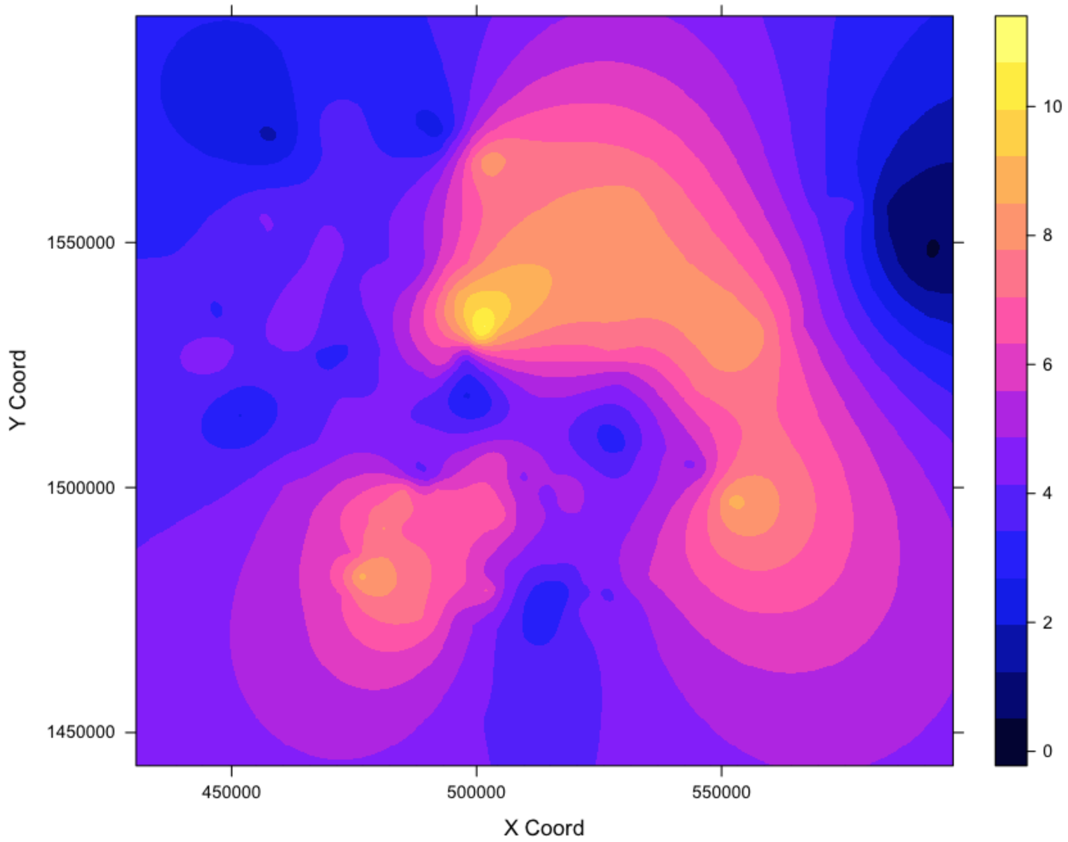



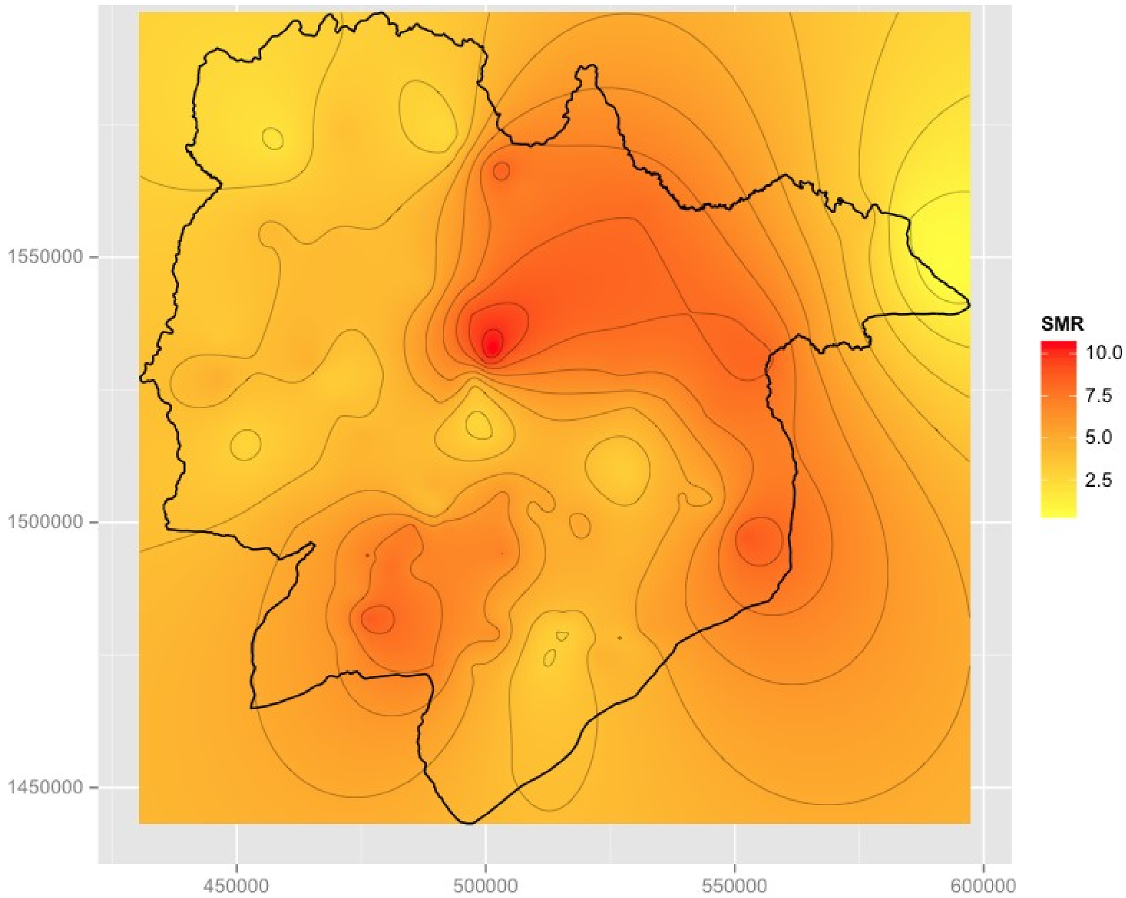
**(c) Preah Vihear, 2010 (IDW)**



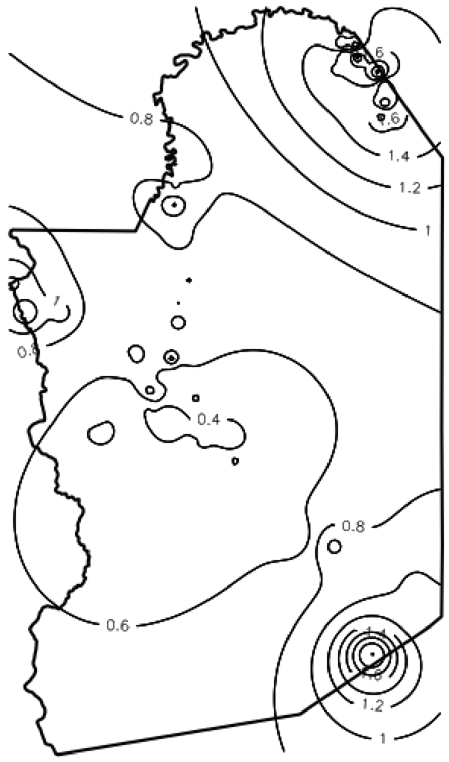
**(d) Preah Vihear, 2010 (Ordinal kriging)**



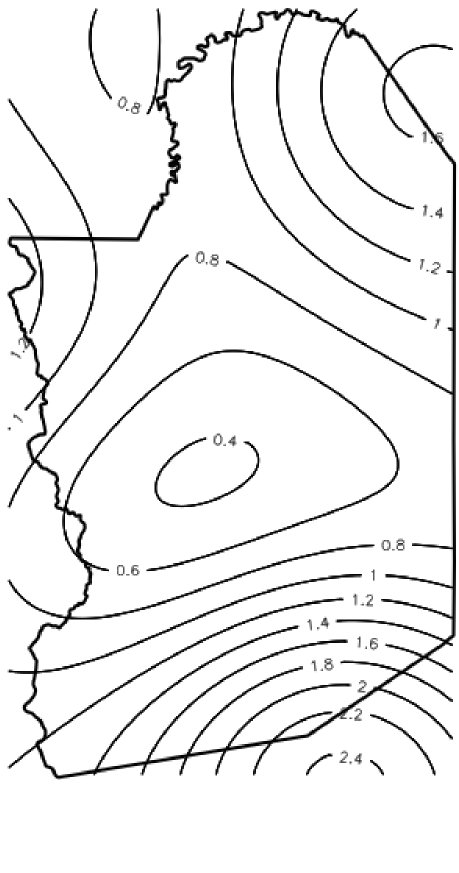



**Risk contour**

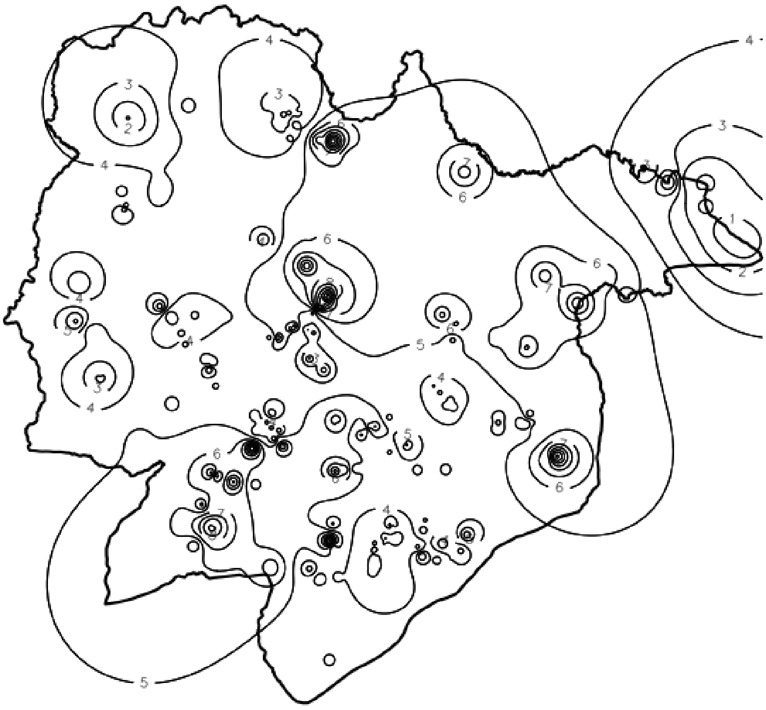
**a**




**b**



c



d

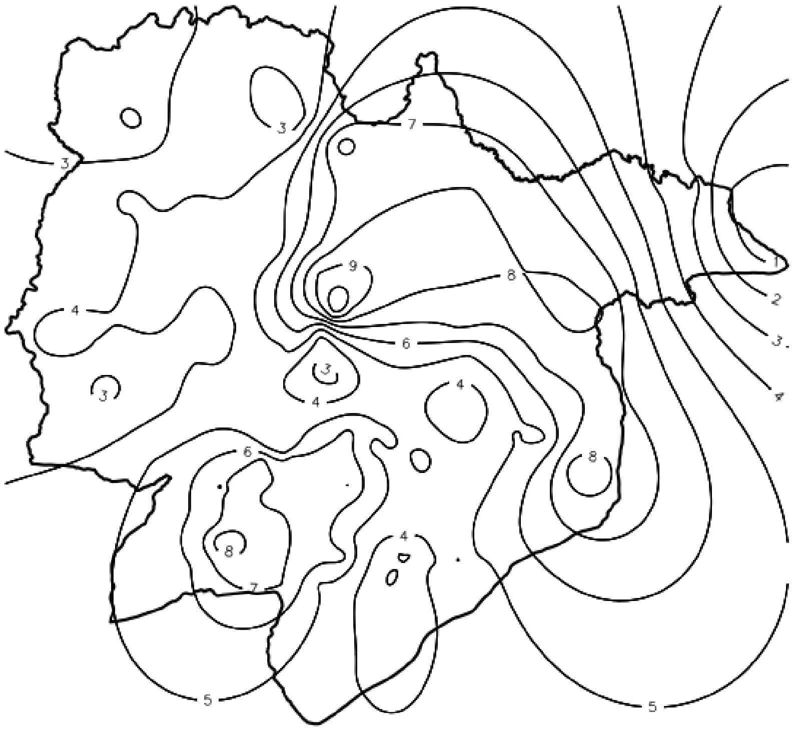
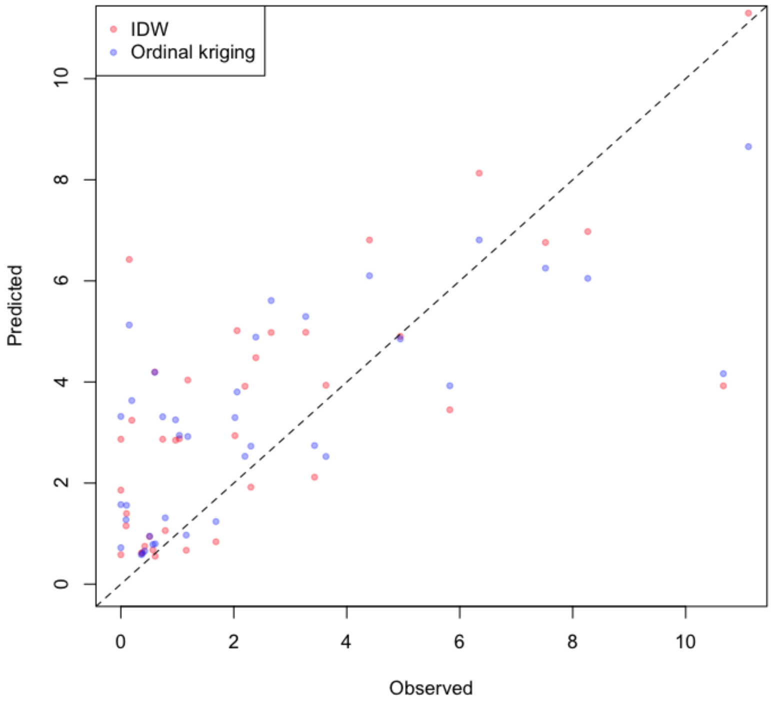


**Figure 20 Representative maps created using the proposed model for Pailin (a, b) and Preah Vihear (c, d) provinces in 2010**

(a, b) Maps of Pailin province and (c, d) for Preah Vihear province in 2010. (a) and (c) were the risk maps created by the inverse distance weighed interpolation method (IDW) and (b) (d) correspond to the maps created by the ordinary kriging.

(a)




(b)

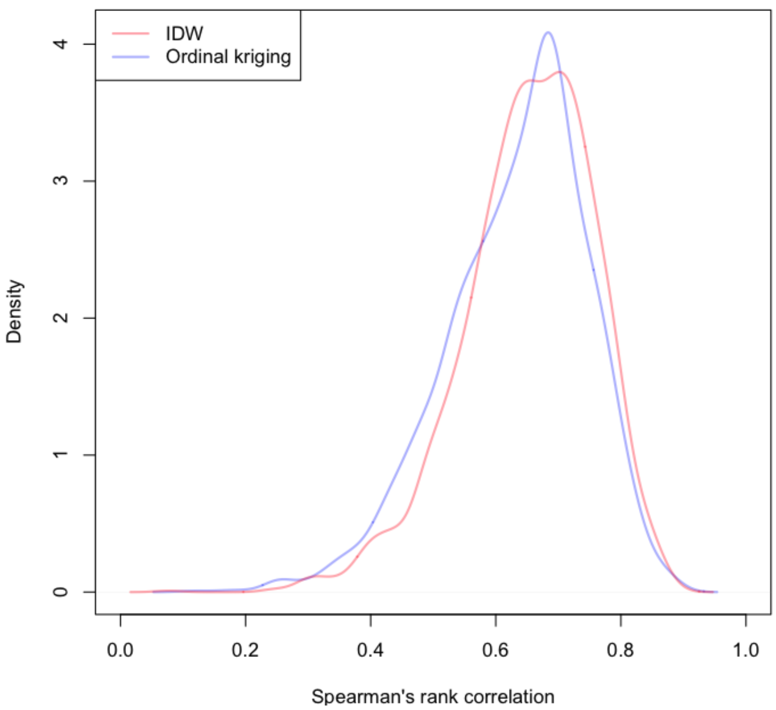
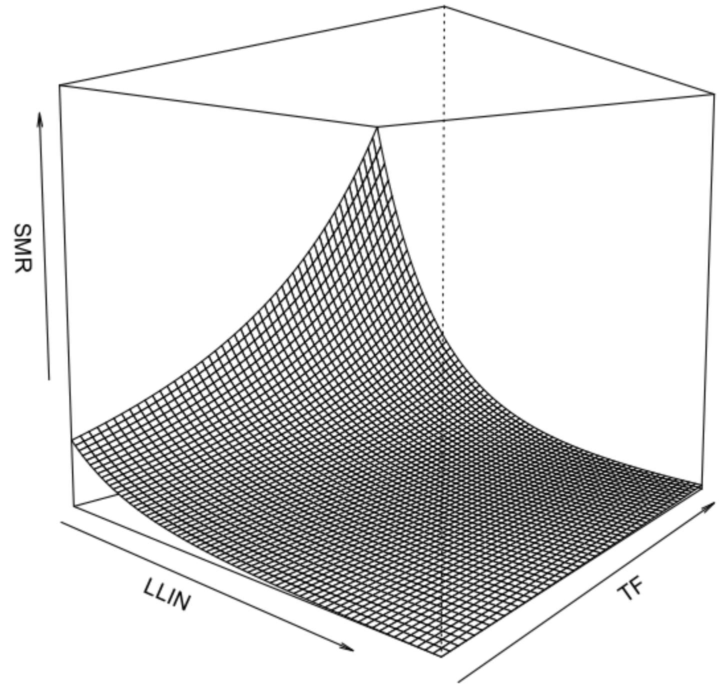



**Figure 21 Comparison of the standardized morbidity ratio calculated from geocoded case data with corresponding predicted values (a) and the kernel density plot of the resampled spearman's rank correlation (b) in the risk map created by the model.**

The dashed line represents 1:1 relationship of observed and predicted values
IDW, Inverse distance weighed method

**(a)**




**(b)**

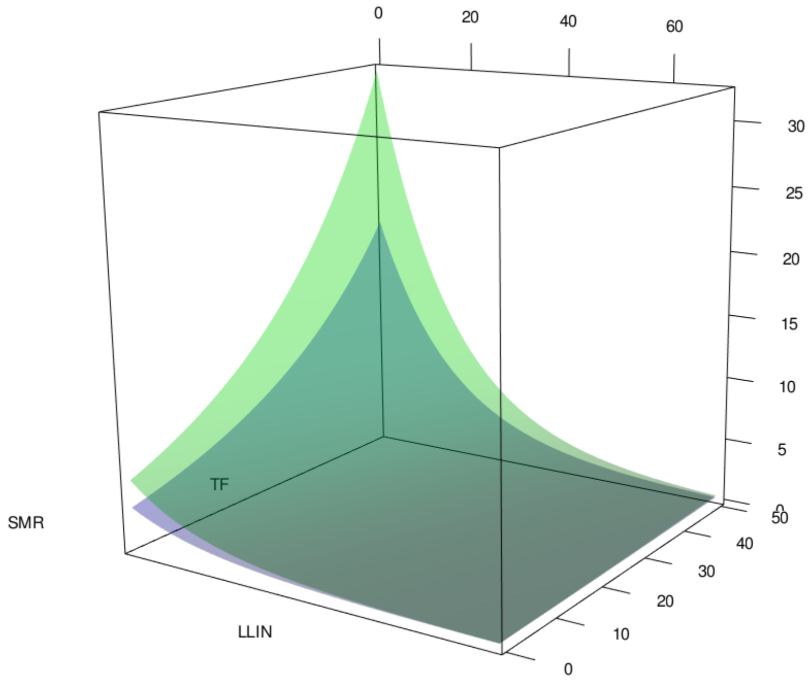
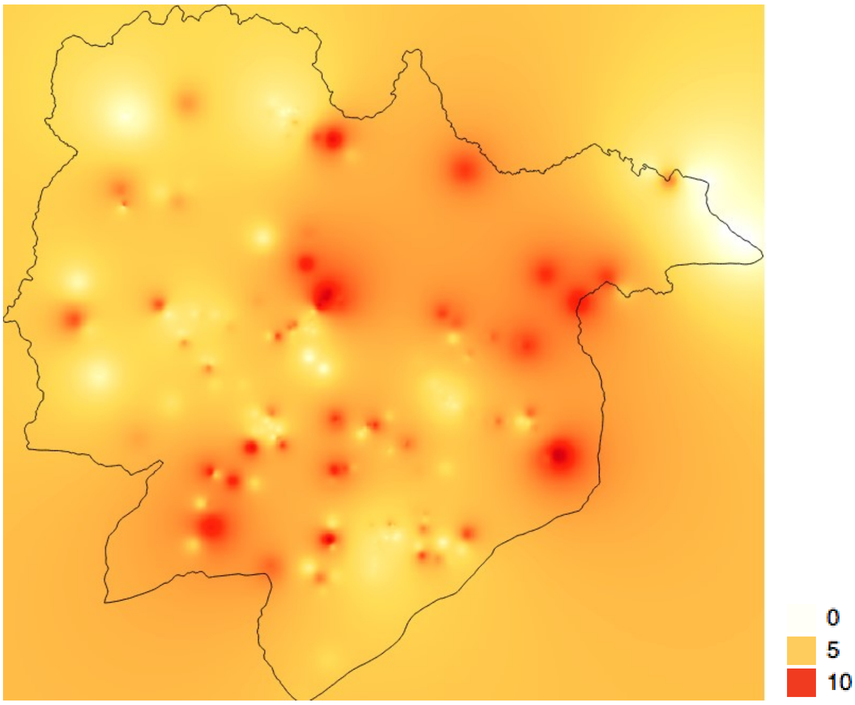


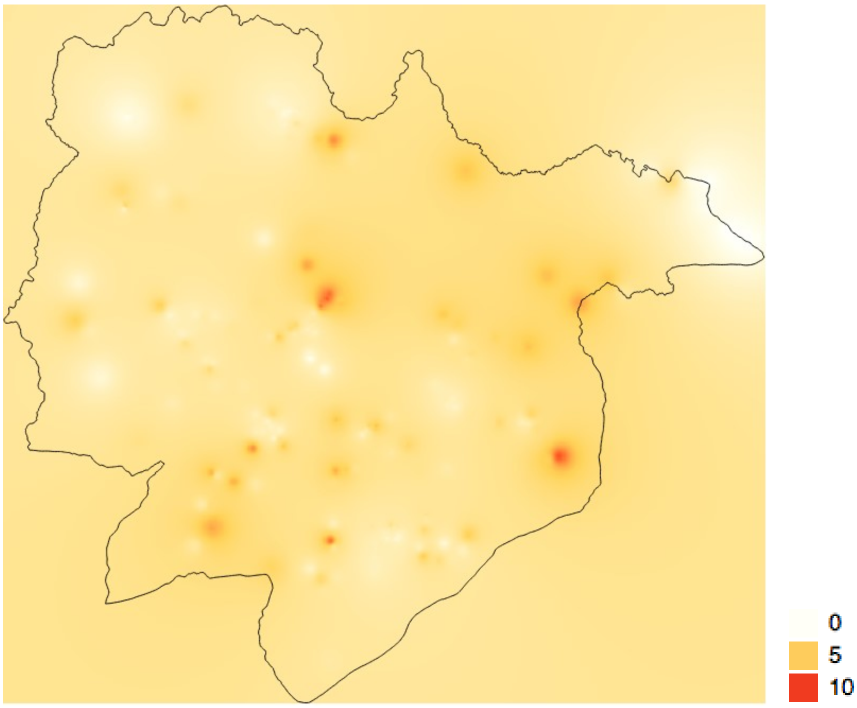

**Figure 22 Computational simulations of expected standardized morbidity ratio (SMR) under various conditions of LLIN coverage and Treatment failure rate of artemisinin**

(a) Relationship of two containment status indicators with expected SMR in Pailin province. (b) Different pattern of expected outcomes from the combination of two containment status indicators in two provinces. The green surface corresponds to Pailiin, and the blue surface corresponds to Preah Vihear province.

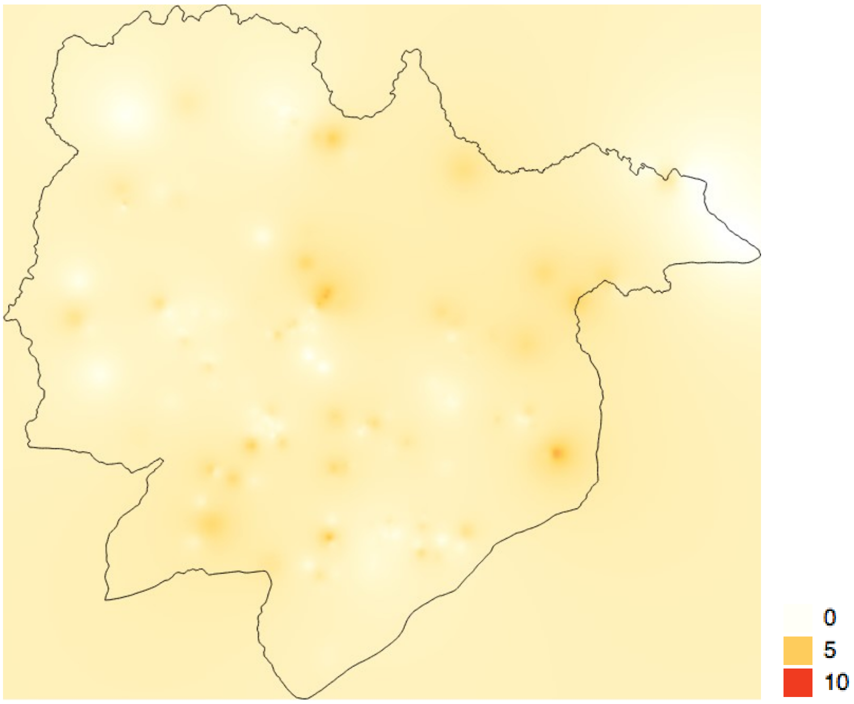
a: Current state




b: TF decreased to 0% (from current 14.7%)



**c: LLIN increased to 20% (from current 5.5%)**




**d**

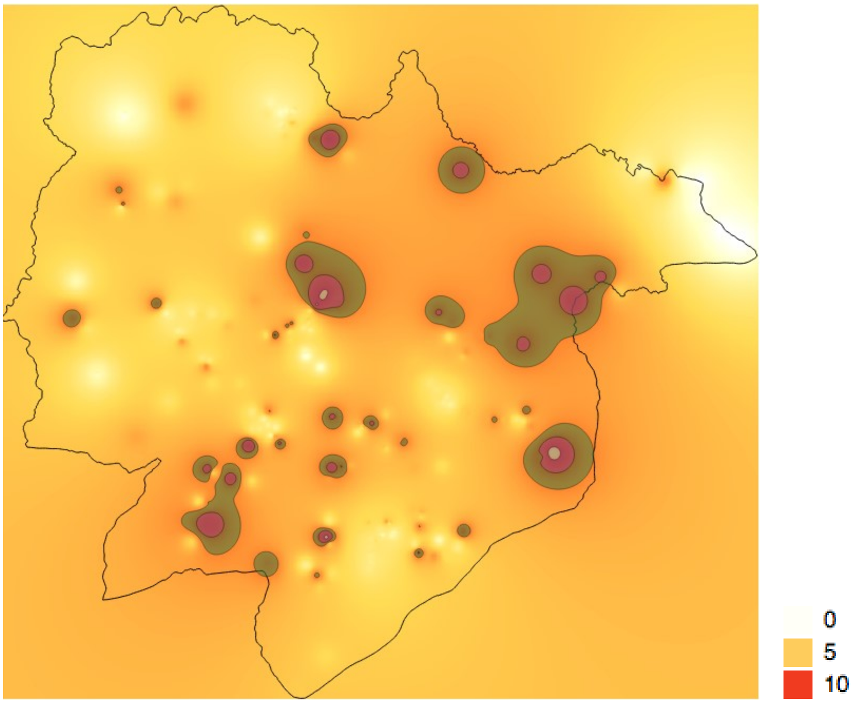



**Figure 23 Case of geographical analysis of expected outcomes from targeted containment status in Preah Vihear**

(a) Current predicted state. (b, c) Mapped geographical view of expected outcomes from targeted containment status. Simulation outcomes can be visualized on the same map (d), by which the effect can be examined. The green (for TF decrease), purple (for LLIN coverage increase) and gray (combined) colored area corresponds to areas at more than 5 SMR.
LLIN, Long lasting insecticide-treated net; TF, Treatment failure rate of artemisinin.

## 3.6 Spatiotemporal modeling and the creation of a fine-scale risk map

### 3.6.1 Spatiotemporal risk modeling

The model developed here successfully explained regional malaria risks. However, the risk of malaria may dynamically alter in accordance with various factors such as the progress of containment interventions and environmental changes. The prevalence or incidence of malaria at given time can be quite variable, due not only to seasonal oscillations but also to complex dynamic factors, including the behaviors of mosquitoes and people, land-cover, housing quality, and the robustness of the health system [100]. To address the complex and dynamic nature of situations within low-to-moderate malaria transmission settings, we build a spatiotemporal model of SMR of malaria incidence. Under the condition that the logarithmic EBSMR ($\hat{\theta}$) follows the Gaussian distribution, the relationship between $\hat{\theta}$ and space-time covariates was modeled using a generalized linear regression model as a function of the *N* predictive variables (*X, Z*). However, given that the situation in respect of malaria transmission may dynamically change in low-to-moderate transmission settings, the model needed to incorporate temporal changes. Furthermore, hidden factors not considered in the model can affect malaria risk and situations may differ depending on areas. To incorporate specific local conditions and temporal changes in the studied areas, we introduced two location or temporal specific parameters, $\varphi$ and $\tau$, to the regression model as in (16) and (17):

$$\hat{\theta} = e^{\lambda} \qquad (16)$$

$$\lambda = \alpha + \sum_{N} \beta_N X_N + \sum_{N} \gamma_N Z_N + \varphi + \tau + \varepsilon \qquad (17)$$

where $\alpha$ is the model intercept, $\beta$ is the parameter associated with environmental covariates *X*, $\gamma$ associated with the non-environmental anthropogenic covariates *Z*, and $\varepsilon$ represents the residual error effects. The location parameter $\varphi$ is the location specific effect that originates from an area's particular conditions, and $\tau$ is the temporal specific effect at every time interval that the data were modeled. We set the time interval as 1 year, and modeled the risk of malaria every year between 2010 and

2013. The MCMC method in the Bayesian modeling frame was employed to estimate the uncertainty about the relationship represented by $\alpha$, $\beta$, $\gamma$, $\varphi$, and $\tau$. The numbers within the estimated uncertainty range of location and temporal specific effects are greatly increased compared with a model which does not incorporate these specific parameters. We employed hierarchical Bayesian modeling to estimate the relatively large number of parameters compared with the amount of data for model building, that is, we introduced a hierarchical prior uniform distribution $dunif(0,\ 10^4)$ for the $\sigma$ of specified non-informative normal prior distribution $N(0,\ \sigma^2)$ of $\varphi$ and $\tau$ for every location and interval of time. For the estimation of parameters of environmental and non-environmental anthropogenic covariates, we specified the non-informative normal distribution with mean zero and large variance, $\sigma = 10^4$. The models were fitted using R software (https://www.r-project.org) on an HOD scale. An MCMC sampler in the JAGS framework, a program for analysis of Bayesian hierarchical models using MCMC simulation [93], was employed for the Bayesian model fitting. Three MCMC chains with 50,000 iterations as burn-in, and 30,000 iterations thinned every 30, were stored as parameter estimates. The convergence of the model was examined using Gelman-Rubin diagnostics [94] and through visual assessment of the trace plots of chains. The estimates of MAE were calculated to quantify the discrepancy between predicted and observed values. Likewise, root mean square errors (RMSE), for assessing the overall model performance, and Pearson's correlation coefficient, were calculated to compare the predicted and observed values at the HOD level. The fitted model was applied to estimate the village level SMR using environmental covariates extracted from the location of each village, in conjunction with specific covariates for each HOD, and at each specific interval of time. The estimated values of the village level SMR were then used as skeletons of the spatial interpolation, using the IDW method. Calculated values, using spatial interpolation methods, were plotted in each 250 x 250 spatial grid at each time interval, from 2010 to 2013, in the two western Cambodian provinces, Pailin and Preah Vihear.

### 3.6.2 Visual presentation of relative weights in the routine surveillance network

In addition to the geographical analyses, we employed graph theory analysis to visualize the estimated priorities of constituent bodies in routine surveillance network, that is, the relative weights of network constituent priorities, based on the spatiotemporal risk model developed. We build a network model of the routine surveillance network in Pailin province using information collected through a survey of published literature, official public documents (guidelines and presentation), and interviews of stakeholders, such as with staff at regional health centers [45, 80, 101]. Health facilities, as well as village malaria workers, report the number of treated cases of malaria patients to higher levels of authority within the surveillance network [80]. Therefore, the focus of this study was on the flow of the reported data for building connections between constituent bodies in the network model. We then multiplied the SMR extracted from the map build in this study with eigenvector centrality values [102] of each network node, as a measure of influence to enhance sensitivity to risk changes at the nodes that can affect to other nodes. The average values of SMR in the surrounding 1-km circular buffers of health facilities were extracted from the risk map in Pailin. The SMRs for the network nodes above the HOD level are considered to be 1, since their networking role is to aggregate reports from health facilities and village malaria workers rather than treat malaria cases. To compare results at different time points, calculated values were normalized to be in the range of 0 – 1, through the min-max normalization method, and then used to represent the size of a network node when network models were plotted at each interval of time.

## 3.7 Results of spatiotemporal modeling for mapping and visualization of the relative weights of network constituent priorities

The parameter estimates of each covariates, as well as their uncertainty ranges, are shown in Table 4. The model showed good convergence, as confirmed using visual assessment of trace plots of chains and Gelman-Rubin diagnostics < 1.01 for all parameters and a deviance information criterion of 176.9. Figure 24 presents the observed versus predicted uncertainty range of the EBSMR in respective HOD at each interval of time. The plot indicates good convergence of observed values through the uncertainty range of predicted values, which presents 87.5% of observed values within the 10th percentile to 90th percentile range, and 98.61% of values covered through a 95% confidence interval of predicted values. The MAE and RSME of the model calculated, using median predicted values and observed values, were 0.328 and 0.626, respectively.

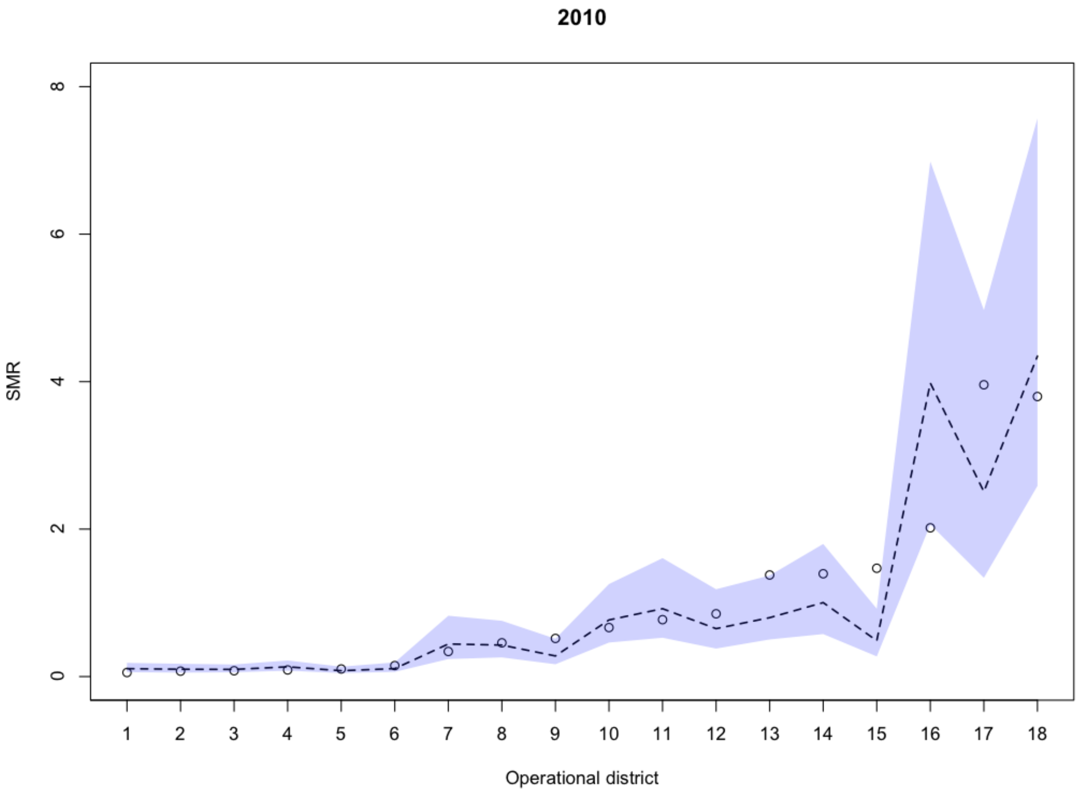
**Table 4. Parameter estimates of covariates of the Bayesian modeling frame and their uncertainty ranges**

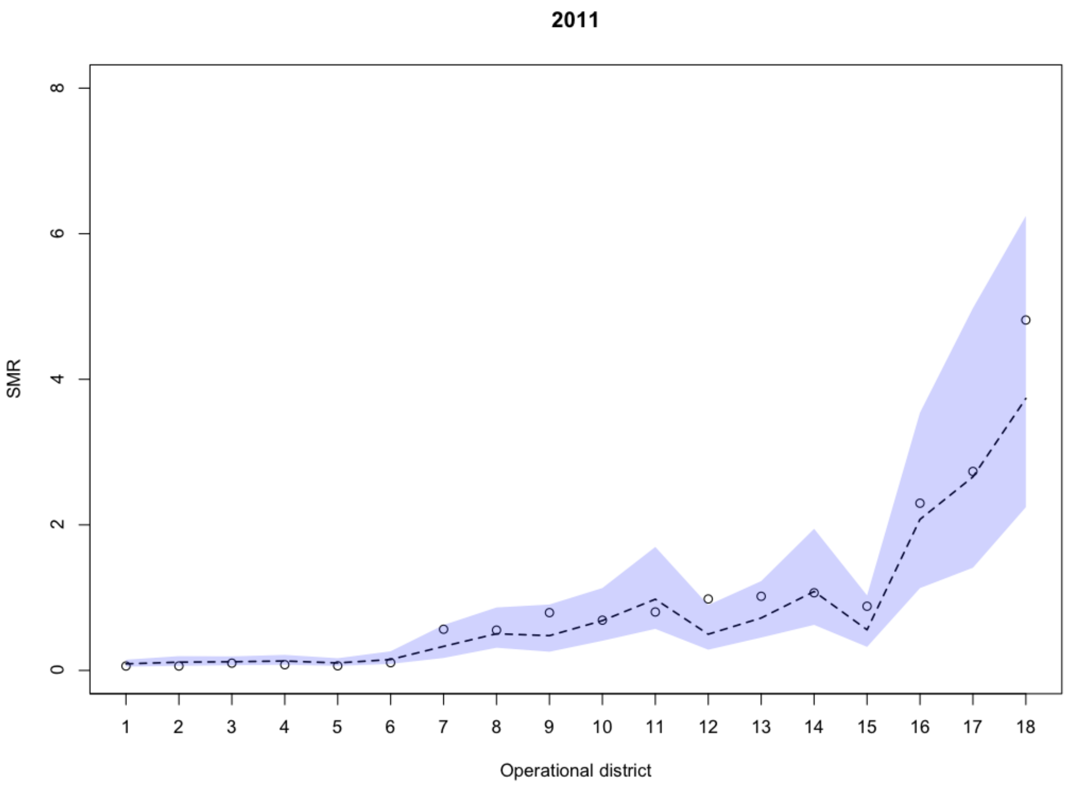
| Parameter | Mean | Standard deviation | 2.5 percentile | 25 percentile | 50 percentile | 75 percentile | 97.5 percentile |
|---|---|---|---|---|---|---|---|
| Intercept | -12.194 | 11.584 | -34.206 | -19.566 | -12.455 | -5.268 | 12.106 |
| NDVI (5 km) | 7.391 | 4.083 | -0.869 | 4.841 | 7.405 | 9.997 | 15.437 |
| NDWI (1 km) | -26.992 | 14.684 | -58.940 | -35.807 | -26.018 | -17.248 | -0.0753 |
| TWI (5 km) | -1.653 | 1.408 | -4.460 | -2.544 | -1.631 | -0.765 | 1.001 |
| *P. falciparum* Temperature suitability index (*Pf*TSI) | 0.00026 | 0.00009 | 0.00008 | 0.00021 | 0.00027 | 0.00032 | 0.00042 |
| Sufficient ownership of LLIN [a] | -0.0632 | 0.0160 | -0.0963 | -0.0730 | -0.0628 | -0.0527 | -0.0331 |
| Treatment failure rate by artemisinin combination Therapy [b] | 0.0377 | 0.0182 | 0.00181 | 0.0264 | 0.0374 | 0.0488 | 0.0754 |

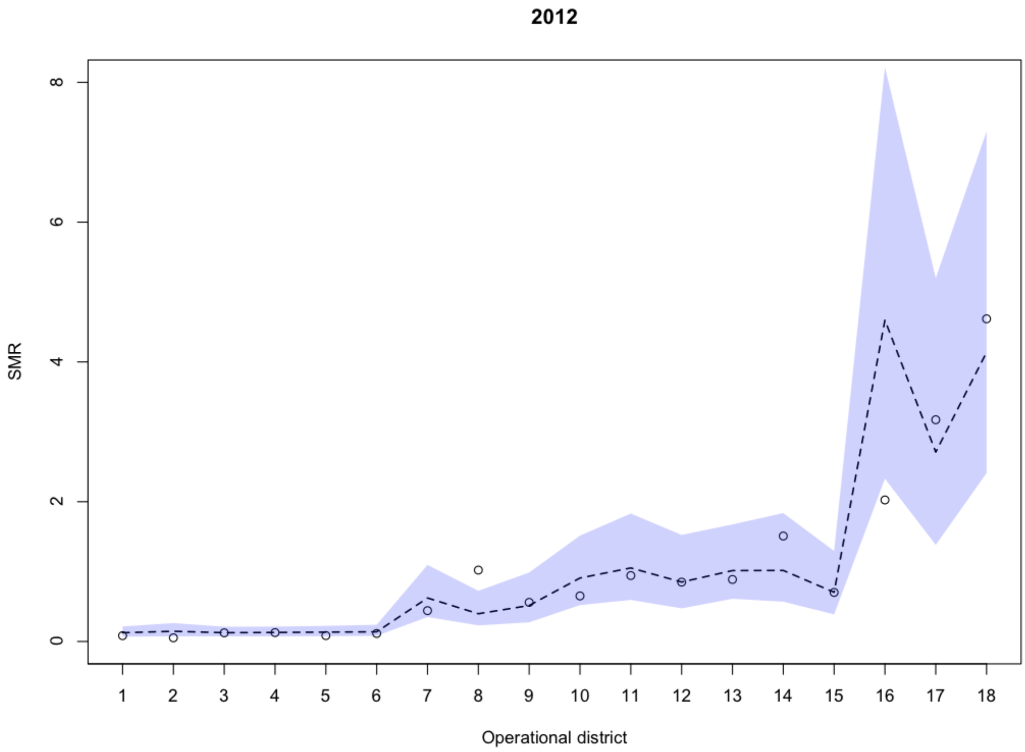
[a] Proportion of households in which distributed mosquito nets cover no more than 2 persons per net.

[b] Positive test for *P. falciparum* on day 28 or day 42

NDVI, Normalized difference vegetation index; NDWI, Normalized difference water index; TWI, Topographical wetness index; LLIN, Long lasting insecticide-treated net


2010
SMR
0
2
4
6
8
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Operational district


2011
SMR
0
2
4
6
8
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Operational district



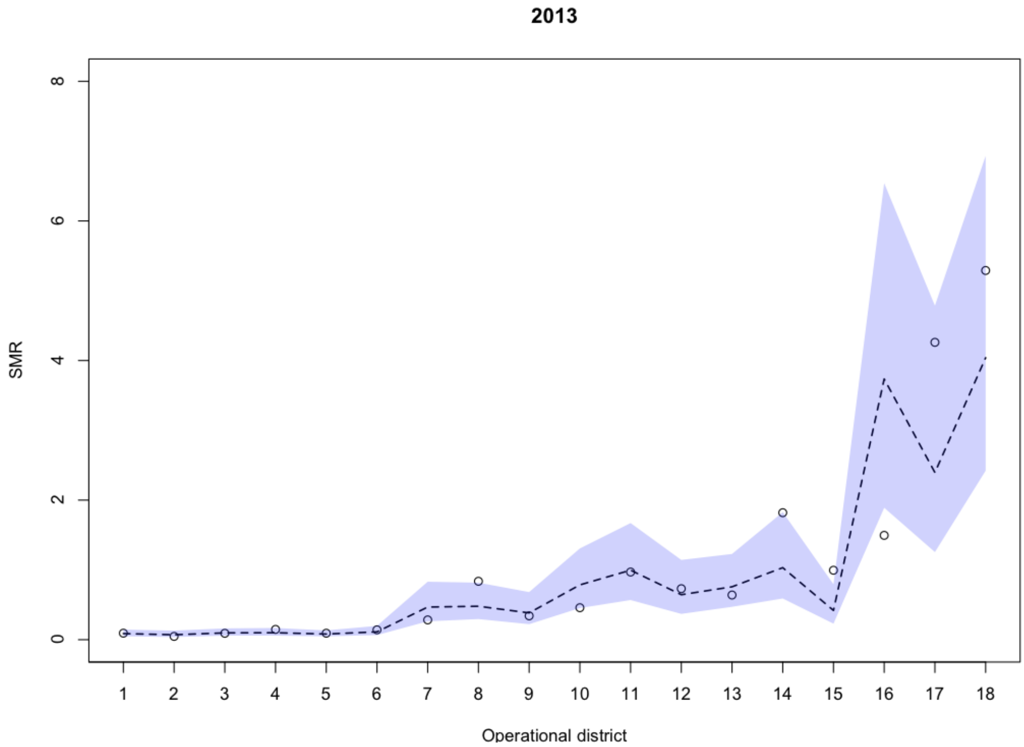
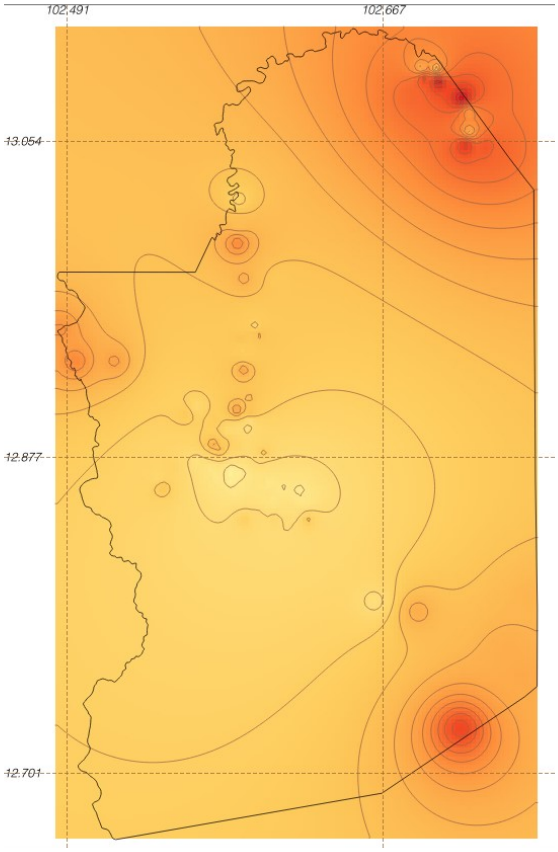




**Figure 24 Observed versus predicted uncertainty range of the SMR in operational districts at each interval of time during the study period (2010 – 2013)**

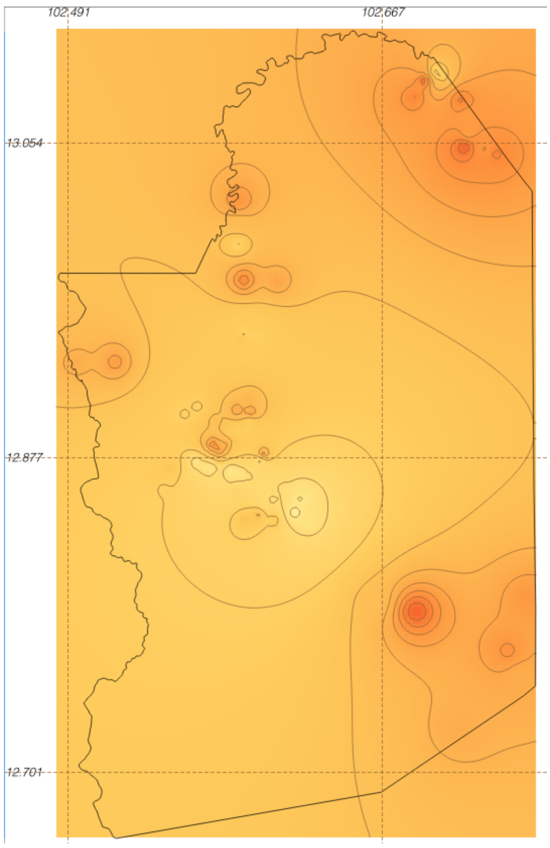
Numbers presented below the horizontal axis indicate respective health operational districts; 1: Sangkae, 2: Preah Net Preah, 3: Thma Koul, 4: Mobkov Borei, 5: Thma Puok, 6: Ou Chrov, 7: Bakan, 8: Sampov Luon, 9: Mong Russei, 10: Siem Reap, 11: Battambang, 12: Sot Nikum, 13: Ankor Chhum, 14: Pailin, 15: Kralanh, 16: Sampov Meas, 17: Samaraong, 18: Tbeng Meanchey

The Pearson's correlation coefficient of observed and predicted median of the EBSMR, using the model, was 0.870 ($p < 0.001$). The estimated SMRs for the villages in the study areas were calculated using the Bayesian modeling framework, and interpolated using the IDW method, to create fine-scale maps of the study area. Figure 25 and 26 show the maps of Pailin and Preah Vihear provinces created by the model developed at each interval of time. As shown in these maps, different patterns of malaria risk distributions at each interval of time are presented. Whereas the place of malaria hotspots did not change dramatically during the study period, the magnitude of risk at these places differed at each interval of time. The visual representations of hotspots in the fine-scale map created here were well aligned with actual areas at high risk, already identified through other sources [90, 96] as well as in another work which was validated by an examination of alignment between the estimated risk and the risk calculated by geocoded case data [103]. These results indicate that the maps created by the present approach do not misguiding presentation.

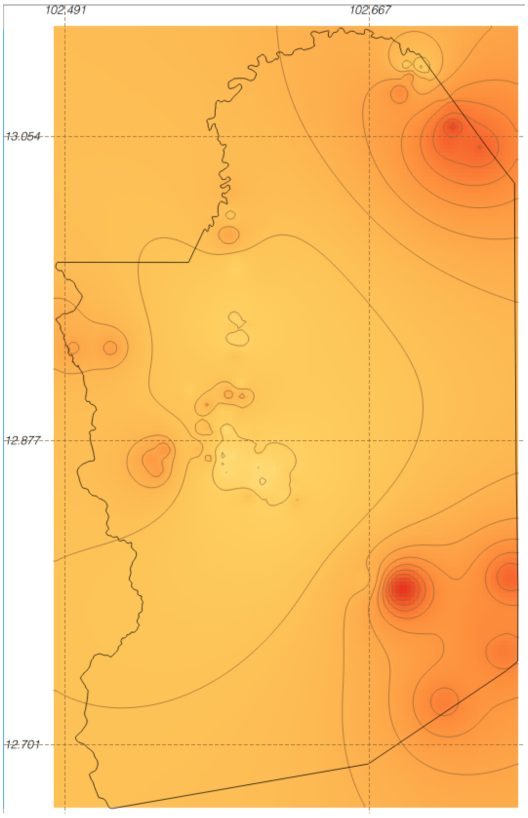
**(a) 2010**


102.491
102.667
13.054
12.877
12.701

**(b) 2011**


102.491
102.667
13.054
12.877
12.701

**(c) 2012**




**(d) 2013**

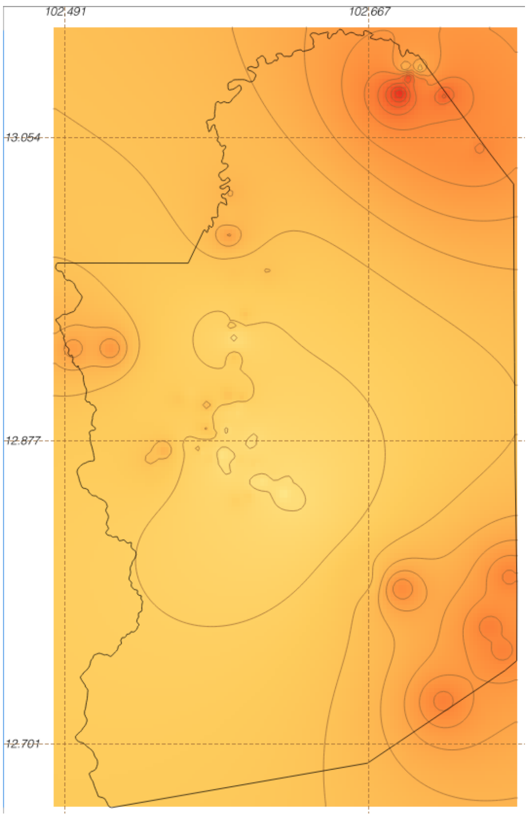
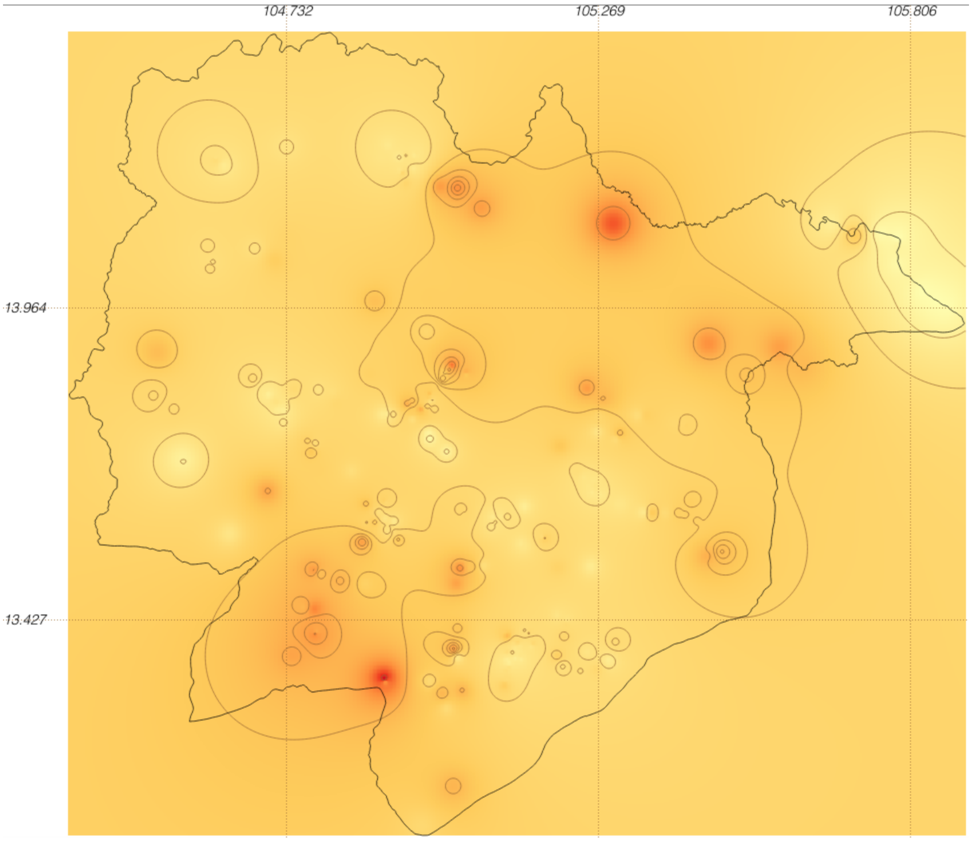
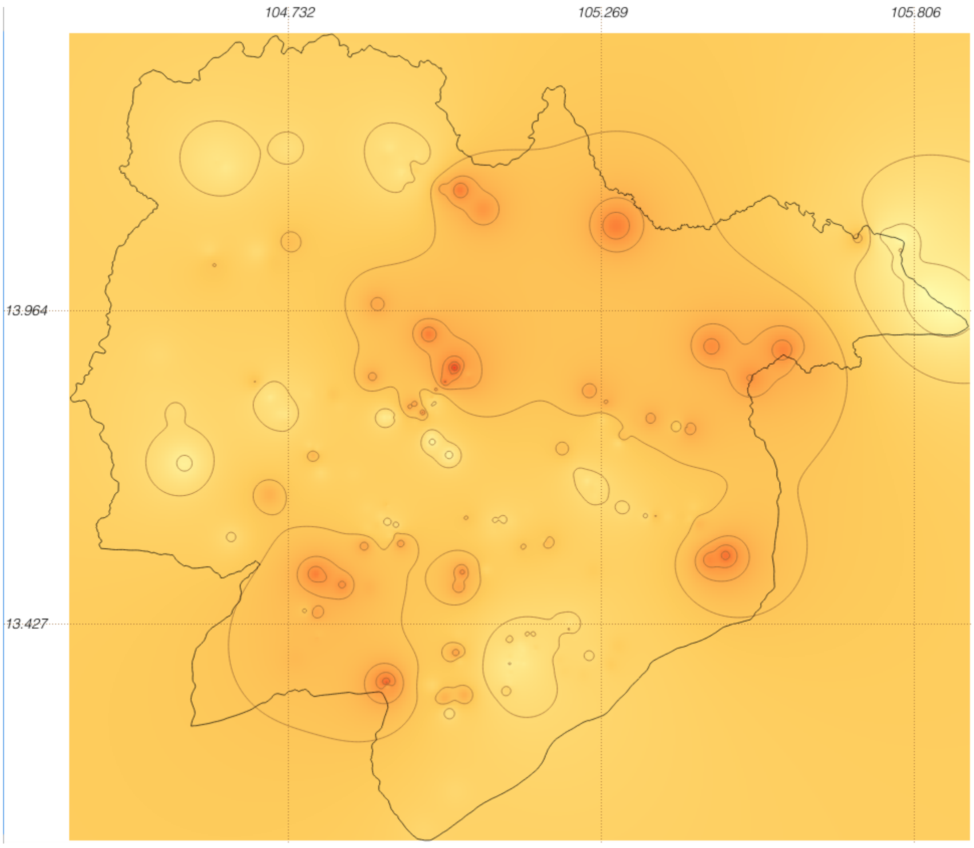



**Figure 25 Maps of Pailin province at each time interval in 2010 – 2013**

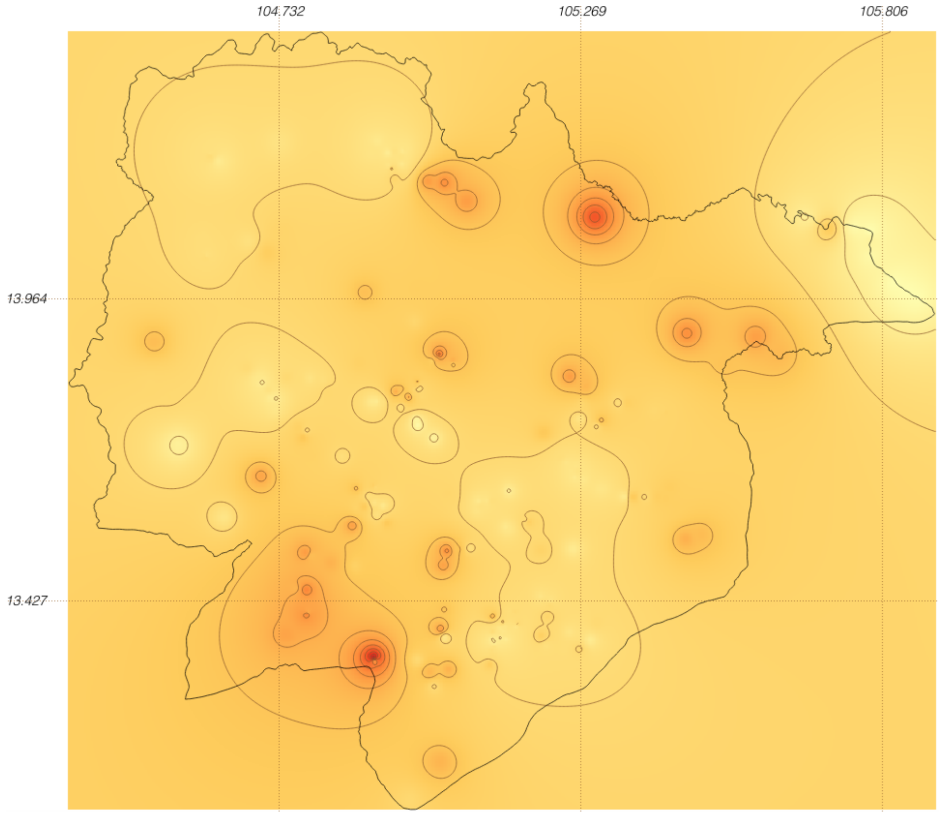
**(a) 2010**


104.732
105.269
105.806
13.964
13.427

**(b) 2011**


104.732
105.269
105.806
13.964
13.427

**(c) 2012**




**(d) 2013**

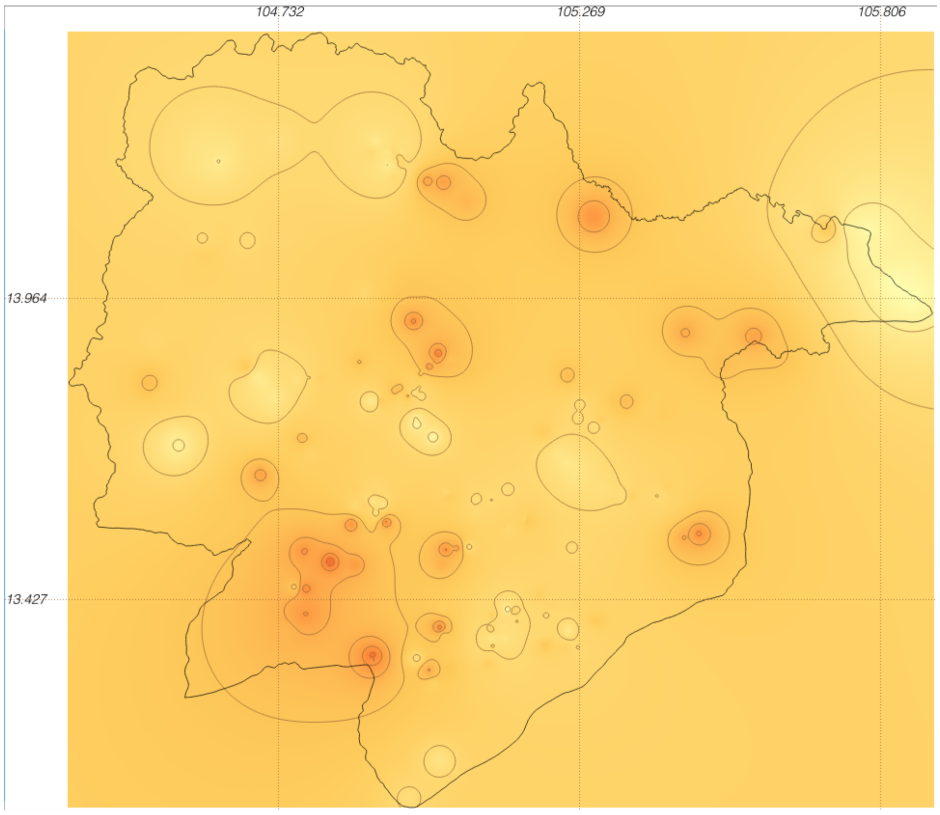
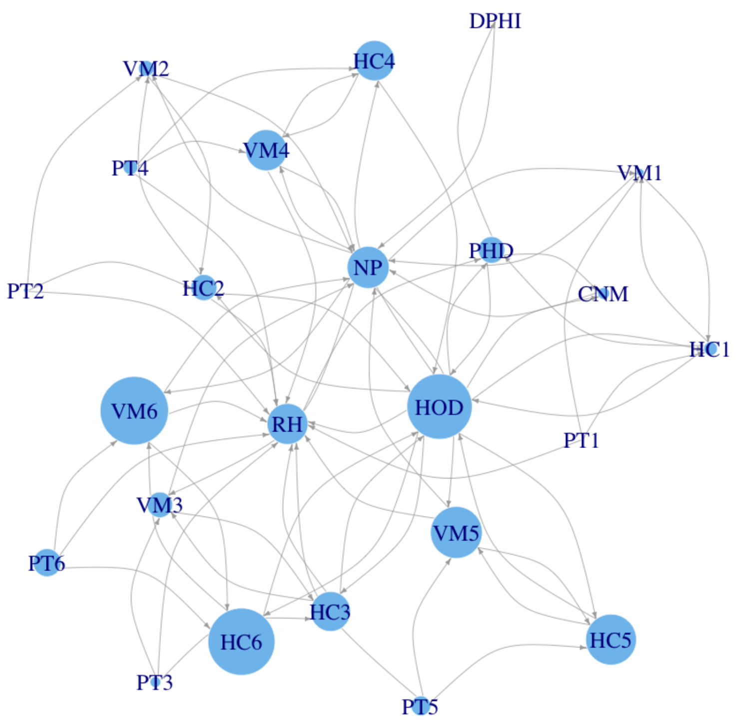


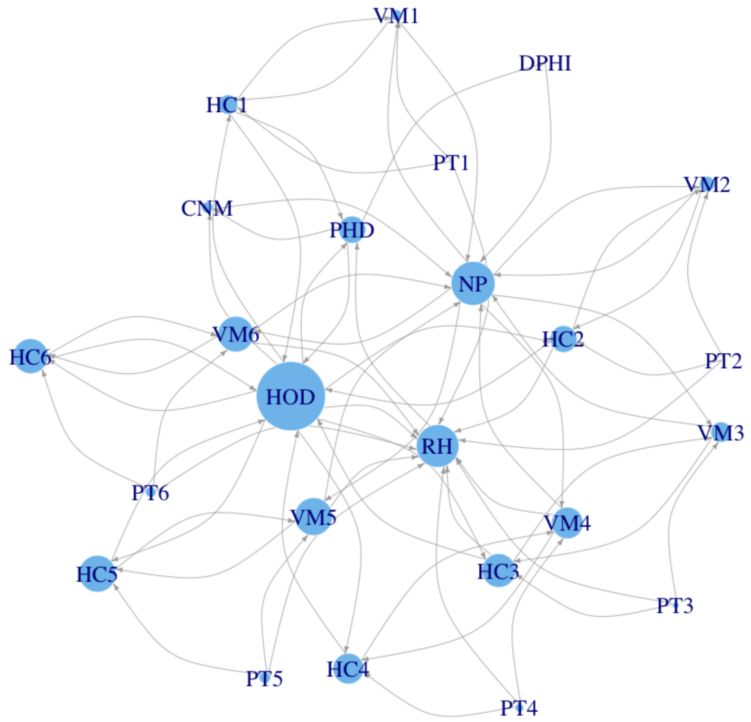

**Figure 26 Maps of Preah Vihear province at each time interval in 2010 – 2013**

Next, the routine malaria surveillance network in Pailin province was modeled, based on collected information and interviews. Figure 27 presents the visualized weights estimated by the risk model, and the structure of the modeled routine surveillance network, at each interval of time. As shown in the visualized network, the different patterns of the relative weights at each interval of time are also presented. In an association with changes in malaria risk in the locations of network constituents on the map, their relative weights in the network were also changed accordingly. Notably, the magnitude of the changes was greater in several peripheral network nodes, such as in health centers (e.g., HC-5 and HC-6) and with village malaria workers (e.g., VMW-5 and VMW-6) than in those of central nodes (e.g., CNM, DPHI, NP, PHD, HOD, and RH). The calculated values of relative weights in the network are listed in Table 5.

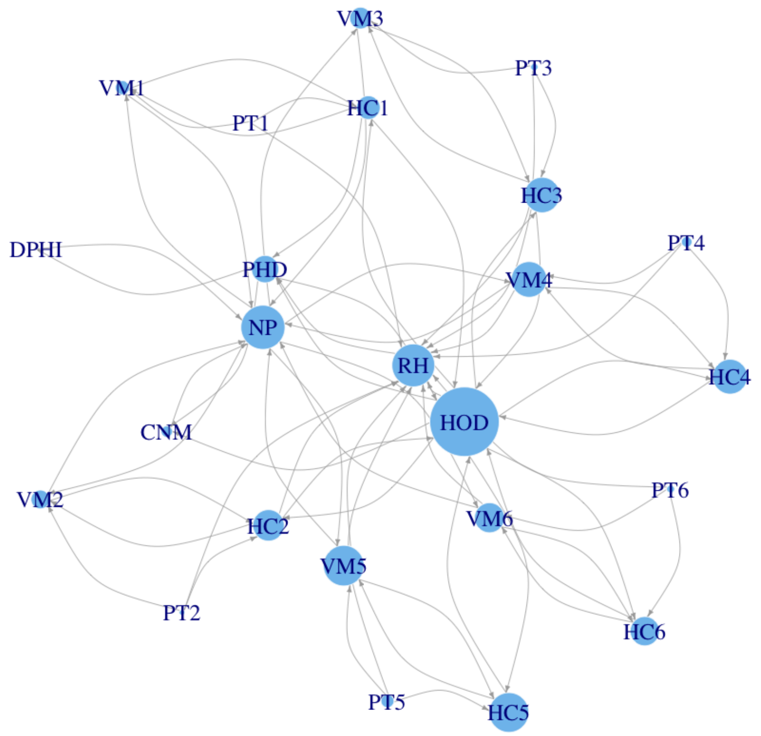
**(a) 2010**



**(b) 2011**


VM1
DPHI
HC1
PT1
VM2
CNM
PHD
NP
VM6
HC2
PT2
HC6
HOD
VM3
RH
PT6
VM5
VM4
HC5
HC3
PT3
PT5
HC4
PT4

**(c) 2012**


VM3
PT3
VM1
HC1
PT1
HC3
PT4
DPHI
PHD
VM4
NP
RH
HC4
HOD
CNM
PT6
VM2
HC2
VM6
VM5
PT2
HC6
PT5
HC5

**(d) 2013**

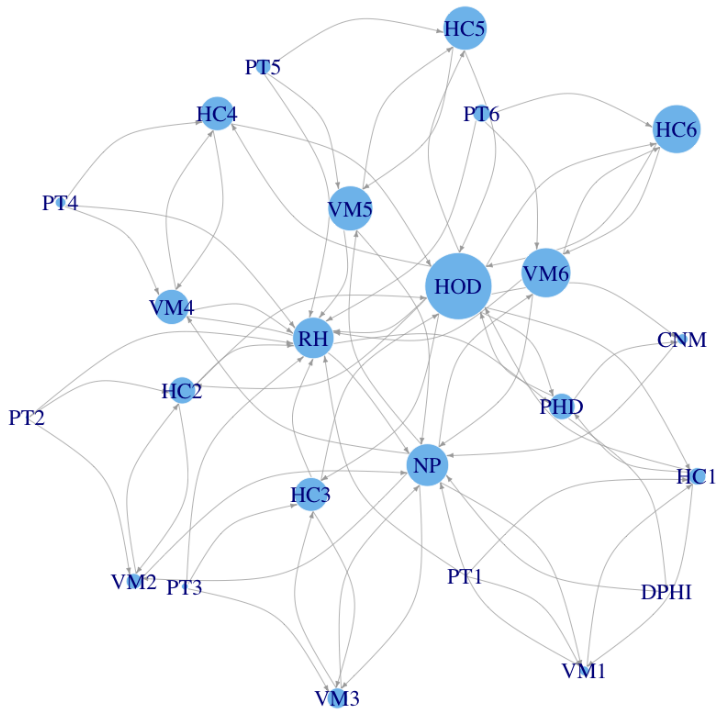



**Figure 27 Visualized weights estimated by the risk model and the structure of the modeled routine surveillance network in Pailin province at each interval of time in 2010 – 2013**

Numbers follow the abbreviations of each network constituent body indicate the subdivided locations of them, including 1: Suon Koma, 2: Ou Chra 3: Phnom Spong, 4: Psar Prum, 5: Phnom Preal, 6: Kracharb. CNM: National Center for Parasitology, Entomology and Malaria Control, NP: National Program, DPHI: Department of Planning and Health Information at the Ministry of Health, PHD: Provincial Health Department, HOD: Health Operational District, RH: Referral hospital, HC: Heath center, VMW: Village malaria worker.

**Table 5. Calculated values of relative weights of key network constituents at each interval of time during study period (2010 – 2013)**

| Parameter | Eigenvector centrality | Calculated values by each year | | | |
|---|---|---|---|---|---|
| | | 2010 | 2011 | 2012 | 2013 |
| CNM | 0.298 | 0.175 | 0.162 | 0.163 | 0.163 |
| NP | 0.694 | 0.615 | 0.635 | 0.635 | 0.635 |
| DPHI | 0.162 | 0.023 | 0 | 0 | 0 |
| PHD | 0.491 | 0.389 | 0.393 | 0.393 | 0.393 |
| HOD | 1 | 0.955 | 1 | 1 | 1 |
| RH | 0.679 | 0.598 | 0.617 | 0.617 | 0.617 |
| HC-1 | 0.461 | 0.200 | 0.280 | 0.339 | 0.258 |
| VM-1 | 0.342 | 0.109 | 0.159 | 0.203 | 0.142 |
| HC-2 | 0.489 | 0.382 | 0.389 | 0.451 | 0.400 |
| VMW-2 | 0.351 | 0.230 | 0.224 | 0.269 | 0.233 |
| HC-3 | 0.489 | 0.582 | 0.489 | 0.509 | 0.503 |
| VMW-3 | 0.351 | 0.373 | 0.296 | 0.311 | 0.306 |
| HC-4 | 0.416 | 0.589 | 0.446 | 0.500 | 0.510 |
| VMW-4 | 0.424 | 0.601 | 0.457 | 0.512 | 0.522 |
| HC-5 | 0.416 | 0.742 | 0.533 | 0.571 | 0.657 |
| VMW-5 | 0.424 | 0.758 | 0.546 | 0.584 | 0.671 |
| HC-6 | 0.416 | 0.981 | 0.504 | 0.421 | 0.728 |
| VMW-6 | 0.424 | 1 | 0.516 | 0.431 | 0.743 |

Numbers following the abbreviations of each network constituent body indicate the subdivided locations of them, including 1: Suon Koma, 2: Ou Chra 3: Phnom Spong, 4: Psar Prum, 5: Phnom Preal, 6: Kracharb.

CNM: National Center for Parasitology, Entomology and Malaria Control, NP: National Program, DPHI: Department of Planning and Health Information at the Ministry of Health, PHD: Provincial Health Department, HOD: Health Operational District, RH: Referral hospital, HC: Heath center, VMW: Village malaria worker.

## 3.8 Stakeholders interview

### 3.8.1 Method

To validate the map created, we conducted series of interviews with key stakeholders. At first, following 3 key questions were asked to the interviewees while showing the process and products of the fine-scale map created here and the map of Preah Vihear province from Malaria Atlas Project database as a reference comparator. Additional comments from stakeholders made in the interview were also recorded.

**Table 6 Key questions for the stakeholders interview**

| No. | Questions |
|---|---|
| 1 | Is this approach useful for supporting the planning of healthcare resource distributions and malaria containment actions? |
| 2 | From visual feature perspective, are the hotspots easily identified? |
| 3 | What kinds of improvements are needed to make this be more valid? Possibilities of other applications? |

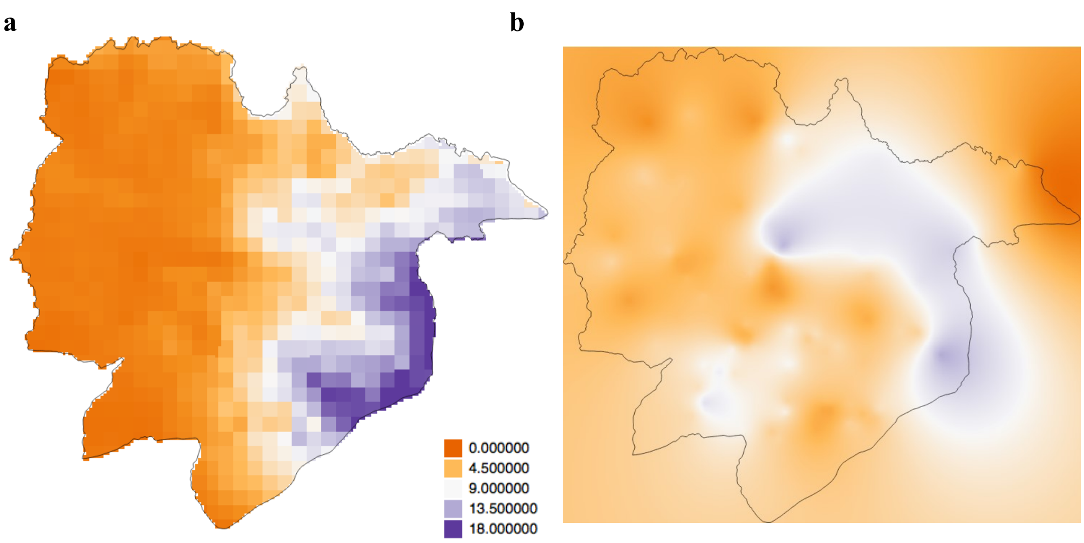



**Figure 28 Visual representations of the map from Malaria Atlas Project database [90] and the fine-scale map created by the risk prediction model developed**

Figure (A) corresponds to the map from Malaria Atlas Project database. Figure (B) is the fine-scale map created by the method developed in this research.

### 3.8.2 Interview with NGO staff

The visual representation of the fine-scale map is very similar to that of Malaria Atlas Project. It makes it easier for the initial selection of interventional target area and its control, which facilitates the impact evaluation by visualize the effectiveness of planned interventions. However, this does not reflect the selective effect on the targeted hotspot. Introducing this map to the staff in the health ministry and the health operational district can be recommended. There are too many reporting steps in the current health information system in Cambodia. The educational level is also related and need consideration.

### 3.8.3 Interview with GIS professional and engineer

This approach is very interesting and the visual representation of the fine-scale map is similar to that of Malaria Atlas Project. Co-kriging can be considered for interpolation method to reflect the regional context appropriately. The most attentive thing is that the map was created mainly from open data sources. However, I think this kind of approach has already done a lot in this research field. Therefore, the uniqueness of this approach needs to be clearly conveyed. The output needs to be simpler, especially for the part of practical application of risk model. Sequence diagram is one option. For country level, this map could give good view. There are many stakeholder acts locally but their action is too detailed. Good customer is the staff of the health ministry. Bill and Melinda Gates foundation could be interested in this product. As for the color presentations, hotspot should be presented as red like colors reminding risk.

### 3.8.4 Interview with field healthcare provider

He thought that the map was useful to see the hotspot and conduct intervention. By 2025, Cambodian government is aiming for eliminating malaria. He thought that this map could support these activities for achieving this target. Most patients come to health center (HC) in January. In August, in the middle of rainy season, people get infected especially in mountain and forest. They regularly go to the mountain side for their work. 24 cases came to the hospital in August. Once patients get treated at HC, they never come back to the HC, unless they get infected again. There are 4 species of malaria, 2 common species are *falciparum* and *vivax*. *P. falciparum* is the most common case in this HC. He reports all the malaria cases come to the HC to HOD.

## 3.9 Discussion

### 3.9.1 Implications from fine-scale maps

As the malaria elimination effort progresses, it has became increasingly important to identify the residual foci of malaria transmission to address the remaining challenges of preventing residual transmissions and preventing the emerging artemisinin-resistant malaria from spreading to protect immunologically susceptible populations. The fine-scale maps that have created here will enable more focused containment actions, such as targeted surveillance, preventive measures, and monitoring for treatment failure, which require intensive support for local health practitioners. A previous report suggested that remarkable proportions of patients in western Cambodia still had parasitemia on day 3 after starting treatment of artemisinin combination therapy, although symptom resolutions were seen within this period [32]. Thus, treatment monitoring is important for preventing patients from discontinuing treatment and developing drug resistance. Interestingly, the visual representations of maps created were similar to those of Malaria Atlas Project; however, maps created in this study displayed a finer level of risk distributions. Some of the differences between the two sets of maps can be explained partially by spatial and temporal variations in the source data. The comparison of predicted risk with geocoded case data confirmed that the areas predicted to be at high risk, which will provide information to quantify expected outcomes from a combination of containment status indictors. These results suggested that these fine-scale maps can play important roles in current situations in Cambodia.

### 3.9.2 Application of SMR for the spatial risk distribution modeling

We also describe an application of SMR using API reported in routine aggregated surveillance data to quantify the spatial distribution of risk by capturing the environmental context and containment status indicators in the model under low-to-moderate transmission settings. We found that the remaining or even increasing tendency of SMR reflected the relative risk of malaria in the studied area during the research period, which can be a useful measure for deciding the allocation of limited healthcare resources. Sturrock et al. built a prediction model using routine aggregated case data and created a fine-scale risk map for Swaziland [63]. In their model, mean temperature and travel time to health facilities were the predictors of both the pixel scale and the coarser district scale of risks. Lowe et al. reported various kinds of predictors such as altitude, living conditions, urbanizations, precipitations and temperature [83]. The variables that we chose for the model were similar in terms of using environmental and human behavior-related variables for malaria risk predictions. Although altitude may be related to malaria ecology, we did not incorporate this variable into the model. Nevertheless, the risk was well explained, probably because of the relatively flat terrain in most of the area that we studied.

### 3.9.3 Environmental context disparities

Of note, the data collection distances from each village for environment-related covariates affected the risk predictions made by the model. The distances selected for the model development for vegetation (NDVI) and water related variables (NDWI), which partially reflects human interactions with the living conditions that exist around human communities. The relationships between *Anopheles* mosquito numbers that cause malaria transmission and distance from mosquito bleeding sites have been reported previously [104-106]. According to surveillance reports [82, 107-108], malaria prevalence decreased by distance from forests. The relationship with the distance from environmental features for malaria risk modeling, such as the proximity of water puddles [109], health facilities [63], have been considered. The effect of distance for the vegetation and water indices used in this study indicates such environmental features are interrelated with human living communities in different ways. Forest workers often working in the forests that are several kilometers away from the communities in which they live, whereas the activity ranges of vectors are limited to short distances from their breeding habitat. The maps created in this study suggest that the spatial heterogeneity of disease risk can be explained by such environmental context disparities. The proposed approach shows that distance from living communities can be a useful reference in which to consider environmental context for cross-scale prediction of disease risk on a fine-scale. The relative risk specified from the surrounding environmental context can be described over a wide area, while maintaining the uniformity of unknown conditions, using remote sensing data by earth observations from space satellites.

### 3.9.4 Implication of computational simulations

It is desirable to use micro data, such as household level data, to build fine-scale risk maps. However, this kind of micro data is often inaccessible and hence they cannot be used for mapping. The encouraging results that we obtained for fine-scale risk prediction in the modeling framework enabled the size of effect to be visualized from different combinations of containment status indicators. The simulation results demonstrated that the predicted outcomes were different under each environmental context, which provides an opportunity for evaluating interventions considering environmental situations in target areas. Moreover, the expected interventional outcomes can be mapped, allowing decision-makers to assess different combinations of interventional approaches considering several constraints such as detailed population characteristics, specific local issues, and resource constraints in a target area. The situation surrounding malaria containment actions differs by area. For instance, the coverage of LLIN is higher in some area where the high malaria prevalence was reported and interventional effort has been made for that. Therefore the incremental cost for improving the conditions of each activity may also be different. Figure 29 shows an example of probability sensitivity analysis using the model developed here. The expected outcomes attained as containment status indicators differ by actions. The decision needs to be made based on the predictions. This example demonstrates the provision of tailored information for the targeted hotspot, by which the decision maker can examine the action alternatives based on the required balance of cost and effectiveness.

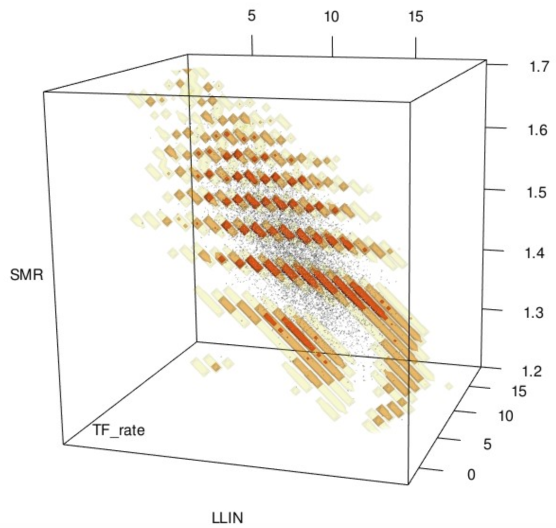



**Figure 29 Example of provability sensitivity analysis under given containment status**

Red-colored grid represents likelihood of expected outcomes under given containment status. Darker color indicates the higher likelihood expected from this simulation.

### 3.9.5 Reliability of data

Generally, the reliability of data is a critical factor for creating relevant models to be used in the real world. The current malaria reporting system in Cambodia relies on the aggregation of field report from village malaria workers and healthcare practitioners to the district, then the province and eventually the national level. Under low-to-moderate transmission settings, passive surveillance systems have difficulty in capturing enough reliable case numbers to reflect the actual situations [110]. Although the variations in the reliability of data reported from each area are likely to exist, the mapping approach described here can add more reciprocity among stakeholders than simply recording aggregated case numbers, which will encourage more effective report-and-utilization cycles and provide an opportunity for effective data utilization. Therefore, this will complement the recent mobile phone based real time case reporting system [80], providing an opportunity for effective data utilization.

### 3.9.6 Spatiotemporal modeling and risk mapping using the developed model

In this study, malaria risks were estimated using information regarding the environmental context surrounding human communities and indicators related to malaria containment, such as the degree of drug resistance and the status of bed net distribution. Not only these global parameters, which can be applied to the whole study area, but also the location and the temporal specific parameters that are used to describe locally and temporally variable malaria risk, were employed to estimate the dynamic nature of malaria risk in low-to-moderate transmission settings. Whereas the visual representations of the maps created in this study were aligned with the original map in which temporal dynamics were not considered [103], changes in the geographical distribution of malaria risk could also be observed between the maps at each interval of time. The greater part of malaria risk factors were explicable through the global covariates, that is, environmental and non-environmental anthropogenic covariates. Like other vector-borne diseases, malaria causation or transmission is highly related to the environmental context surrounding human communities. Remote sensing data captured by space satellites is supposed to be cost-effective in monitoring ground conditions over widespread areas. Furthermore, an opportunity to improve the risk model is available, through accumulating and fitting these data together with malaria case data collected through the routine surveillance system, in an iterative manner. The results also indicate the significance of location and temporal parameters in assisting the ongoing malaria containment effort in low-to-moderate transmission settings. In this study, we employed hierarchical Bayesian modeling frame to incorporate location and temporal specific effects into the risk model. This approach was effective in estimating the uncertainty ranges of a relatively large number of parameters compared with the amount of data. This presentation of the estimated uncertainty range of location and temporal specific parameters gave expression of the spatiotemporal dynamics of malaria risks to identify changing malaria hotspots over time in low-to-moderate malaria transmission settings. Furthermore, the model discriminates among the effects of global parameters, that is, the effects of environmental and non-environmental anthropogenic covariates commonly observed through the study area, and location and temporal specific parameters. This possibility allows for the association of environmental or non-environmental anthropogenic factors with malaria risks to be predicted, with their uncertainty ranges, in nearby areas or on other geographical scales without the bias of each specific effect. While this approach successfully demonstrated the different patterns of malaria risk distributions patterns of malaria at each time interval, it is also important to validate this mapping approach for predictive analyses. One possible solution could be introducing temporal effect modeled by an autoregressive process instead of estimating temporal specific effect. Because of the limited amount of available data, we could not split the data for cross-validation. Instead, we build the model using all data to investigate whether the spatiotemporal analysis can capture the small difference of risk distribution patterns at each time interval. Based on the results of this study, the next step can expand this model for predictive analyses using sufficient amount of data.

### 3.9.7 Transitional complexities exist in routine malaria surveillance network

Complexities caused by the dynamics of malaria endemicity were enhanced through the visualization of weights estimated by the risk model and of the structure of the routine surveillance network. In this network model, the size of the network nodes represented the relative weights scored using the centricity value as a measure of influence, and using the SMR as a measure of relative risk of malaria in the studied area. These measurements can support decision-making around allocation planning of limited healthcare resources in low-to-moderate malaria transmission settings, based on predicted malaria risk factors and the importance of the network constituents. The calculated score indicated that relative weights were undergoing change among several network bodies, for example, among regional HCs and VMWs, whereas scores were stable at some other network constituent bodies, such as those in a more centralized part of the network structure. This observation is likely to become subject to even more complication with changes in the network structure occurring over the course of malaria containment. A study in Africa reported that the loss rate of insecticide treated nets were faster than estimated based on the previous prediction models [111]. This finding indicates the needs for the continuous tracking of regional containment status. Moreover, the study showed that resource distribution inefficiency was caused due to several factors, such as an over-allocation of mosquito nets, which is commonly observed in malaria containment systems. Over-allocation is likely to be a result of a complex web of factors, such as multiple healthcare resource distribution strategies and varying degree of population access to services. The resource requirements at each healthcare facility are likely to be changing over time and influenced by multiple factors. Establishing a continuous feedback cycle of data collection through the surveillance network, and utilization of data for optimal resource allocation planning, while strengthening the system to improve data collection, could be a possible solution to overcome the transitional complexities of the system.

### 3.9.8 Utilization of data from publicly available sources

In this study, almost all data used to build the risk model were publicly available data. This approach provides particular advantages in respect of routine operational costs, provided that data reliability is maintained to a high level. The reliability of data is often a matter of concern in many real-world situations. As reported in a previous study, the quality of data from health facilities may vary, due to various factors. However, the approach described in this study can be used not only to identify target hotspots but also to enable more timely feedback and facilitate more information sharing among healthcare practitioners. This outcome will encourage more effective report-and-utilization cycles and eventually provide an opportunity to improve the quality of care and collected data throughout the entire system as it works toward the malaria elimination.

### 3.9.9 Limitations

While the proposed approach generated several supportive results in terms of fine-scale risk predictions under a low-to-moderate transmission setting, several important limitations and considerations for the future work should be considered. First, containment status indicators other than $LLIN_{suf}$ and $TF_{rate}$ were not considered in the present model. The expected outcomes of interventional efforts could be obtained from the results of various activities, which may not be explained by a simple additive effect, but rather through the interaction of these activities. In the model, we considered the interaction between $LLIN_{suf}$ and $TF_{rate}$, but the result did not improve. Therefore, interactions to describe the complex reality should be considered for practical applications for assessing the effectiveness of interventions. Second, the influence of migrant population was not considered in the modeling framework. The dynamics of human carriers that drive parasite transportation between regions can be quantified using spatially explicit mobile phone data and malaria prevalence information [6]. By incorporating these factors into the modeling framework, more useful models could be developed. Third, to ensure as many cases as possible in the risk modeling, we included cases with both *P.falciparum* and *P.vivax* in the analysis. Considering that treatment differs between these malaria species [112], it is more appropriate to identify discrete spatial and temporal patterns of different malaria species in the analysis. However, it is increasingly difficult to assemble the necessary number of cases to build rigorous risk models of target locations when reported cases become rare due to progress in malaria elimination. A study in Bangladesh reported similar patterns in the association between environmental covariates and the incidence of these two malaria species in the discrete analysis [113]. Given the situation, we consider our analysis provides useful information in low-to-moderate malaria transmission settings. However, complementary data, such as the past trend of malaria incidence arising through different malaria species may be required for more appropriate healthcare resource planning. It is possible to conduct separate analysis for both species by accumulating sufficient case data, which could lead to more appropriate allocations of required healthcare resources to different hotspots to avoid waste. As such, the separate analysis using this approach remains to be confirmed. Fourth, due to their limited availability from publicly available data sources, those covariates related to containment status indicators have not been considered as time dependent. As a matter of course, malaria containment interventions can change along with their progress. Containment interventions are affected due to various factors such as the baseline level of malaria and proportion of people who have already been covered by the containment. As a case in point, a reduction in malaria can be different in respective areas with different baseline levels of malaria if identical interventions are implemented across the entire region [100]. Continuous monitoring of the entire region is important to measure the effectiveness of containment interventions, but can sometimes be costly. The mathematical modeling approach for predicting the effectiveness of containment interventions can be a good alternative for intensive surveillance monitoring in certain situations and may improve the present approach through projecting situational changes. Fifth, we employed the IDW method to study the changes in patterns of malaria hotspots based on the findings that the maps created by this method presented more spotted malaria risk compared with the other interpolation method [103]. However, the interpolation method is depending on several factors such as spatial density of villages. As such, it is more appropriate to evaluate the changing risk patterns using maps created by multiple interpolation methods. Sixth, the treatment seeking behavior varies spatially, which may affect the reporting bias of case data. Sturrock et al. addressed this issue in their modeling approach using Swaziland malaria information system [63]. Unfortunately, this kind of information in Cambodia was not available from publicly available sources. Thus, we need to conduct field survey in the sampled place of target area if this aspect needs to be incorporated. However, in addition to the case reported from public facilities, cases reported from VMWs providing primary healthcare services to the community were also counted in the surveillance report we used. Since the VMW program is active in northwest Cambodia, this structure can improve the coverage of potentially detectable cases to a certain extent. One of the strengths of the approach used in this study is that the maps were created mostly from publicly available data. Therefore, map authors need to collect complementary data from the field if it is necessary considering the balance of timeliness and reliability of the map. Finally, the network model developed in this study was only based on information from available sources and interviews. Hence, we could not fully account for the possible influence of subjective factors in building the model nor for differences arising from unreported situations. Using public available sources makes it is possible to continuously reiterate the model development of both malaria risk and of the healthcare network models and to assess the ongoing situation, without significant cost constraints.

### 3.9.10 Future prospects

Like all programs, malaria elimination action programs need specific plans with realistic time limits and well-defined parasitological and entomological goals [68]. Maps created by the modeling framework here can provide opportunities for establishing realistic goals using current resources. Furthermore, the maps can provide useful information both quantitatively and qualitatively for monitoring and evaluating interventional activities, while providing decision-makers with a platform for cross-scale wandering to help make decisions for efficient healthcare resource use. The proposed approach is simply a quantitative prediction technique for using existing dataset, and thus can only play a part in the whole healthcare information system for malaria elimination. Clearly, the divergences of the prediction from real world situation need to be considered. Nevertheless, the adjustments in malaria quantification contribute further steps in a system that is working toward malaria elimination. Through continuous improvement cycles of the malaria risk model, and through appropriate revisions of healthcare system modeling with the help of various stakeholders involved in the healthcare system, opportunities for optimizing healthcare resource allocation planning in an adaptive manner are likely to be generated, which could contribute specifically to further progress toward malaria elimination.

## 3.10 Chapter summary

We demonstrated a mathematical modeling approach for identifying regional malaria risks, using routine aggregated surveillance report combined with environmental data and non-environmental anthropogenic data. We have demonstrated a mathematical modeling approach for SMR using API from routine aggregated surveillance report and generated cross-scale predictions within a modeling framework that correspond to environmental context disparities to create malaria risk maps on a fine scale. Using the mathematical mode developed, different representations of simulated outcomes from containment status indicators are presented, which provides useful insights for tailored planning of action alternatives considering regional malaria endemicity. A hierarchical Bayesian framework was employed to fit the transitioning malaria risk data onto a map. The model was fitted to estimate the SMRs of every study location at specific time intervals within an uncertainty range. Using a spatial interpolation of the estimated SMR at village level, we created fine-scale maps of two provinces in western Cambodia at specific time intervals. The maps successfully represented different patterns of malaria risk distributions at specific time intervals. Moreover, the visualized weights estimated using the risk model, and the structure of the routine surveillance network, represent the transitional complexities emerging from ever-changing regional endemic situations.

# 4. MANAGING TRANSITIONAL COMPLEXITIES IN THE HEALTH INFORMATION SYSTEM

In this chapter, we demonstrate the approach of modeling and analyzing transitional complexities in the health information system-of-systems. A health information system is a SoS that serves as a foundation of health information used for the evidence-based decision making of public health actions. It also act as foundation of supply source of health information that is used to create risk maps for resource allocation. The quality and timeliness of health information continue to be of concern. Hence, it is important to understand the condition of this SoS in an ongoing manner of effective system engineering. Practical solutions for addressing system transformation and its effect on system management are required given that the SoS transforms over time. As a case, along with the increase in the mobile phone usage and increasing coverage of VMWs, the malaria surveillance needs to collect both the information from health facilities and VMWs, i.e., the data sources are expanded [80]. Accordingly, changes are occurred in the architecture and the process. In such situation, it is required to understand the condition of the SoS for effective use of limited resources. The application of the graph theory for SoS engineering has been extensively discussed in recent years [76]. The knowledge and approaches generated from this attempt can help explore ideas on optimizing SoS and calculating their complexity. Therefore, we studied an approach to model and analyze the transforming health information SoS using the process, architecture, and the risk associated with environment in the case of Cambodian malaria surveillance system. The next section explains the overview of the Cambodian malaria surveillance system and discusses its reformation along with changing environmental and stakeholder needs in detail. Next, the models of the system under investigation are presented using the engineering systems multiple-domain matrix (ES-MDM) modeling framework [114]. A scoring approach is then demonstrated using attributes of the process, architecture, and risk associated with environment to analyze the relative weights of the constituent systems. By comparing the calculated scores with the results of the simulation test employing agent-based modeling (ABM) method, further insights into the progress, the limitations, and advantages of the scoring approach are discussed in the next section. Finally, this study is summarized, and the direction of future work are outlined. To the best knowledge of the author, this approach provides the first step of addressing the needs for continuous management of transforming health information SoS by analyzing the transitional conditions of the constituent systems as well as the whole health information SoS.

## 4.1 Description of the system of interest

### 4.2.1 Overview of the Cambodian health information system

Under the guidance of the MoH, the public health service is delivered through a vertical structure in Cambodia. The MoH oversees health-service delivery through 24 PHDs comprising 81 HODs. Each PHD operates provincial hospitals and governs the HODs. Each HOD covers roughly 100,000 – 200,000 people with RHs primarily delivering secondary care and a number of HCs. HCs cover 10,000 – 20,000 people and provide a minimum amount of care, comprising largely preventive and basic curative services [115]. The national health information system contains information on day-to-day health services and health problems encountered at all levels of health facilities [44]. The current primary components and standardized monthly reporting forms of the system are as follows: HC1 form for the HCs, HO2 for the RH, DO3 for the aggregation of HC1 and HO2 within the HODs, and PRO4 for the aggregation of DO3 within the PHDs, which is sent to the MoH after compilation as the last form [45]. The data collected at the HCs and RHs are aggregated at the HODs. After the compilation of the data at the HODs, they are forwarded to PHDs and the MoH.

### 4.2.2 Reformation of malaria surveillance system in Cambodia

Recent malaria containment activities have decreased the occurrence of malaria in Cambodia by less than half of that in early 2000s [26]. Currently, approximately half the Cambodian population is living in malaria-free or low-transmission conditions [27]. Despite these facts, several issues, such as emerging artemisinin resistance [28, 32] and remaining foci of malaria transmission, need to be solved for the elimination of malaria. As malaria is one of the prioritized public health issues in Cambodia, special arrangements have been made in the routine malaria surveillance. In addition to the usual case reporting, there is a malaria-specific laboratory section in the standardized reporting form, slides, and diagnostic testing kit [45]. Conventionally, the main source of malaria case data is the national health information system, which provides aggregated data at the HOD level. Although these data were useful in obtaining the overview conditions of broad areas, the data requirements for routine surveillance activities are subject to change with respect to those in the finer scale (e.g., village level), including case, demographic, and containment status, such as bed net distribution, to identify the remaining malaria hotspots. In addition, considerable resource constraints in funding and manpower in the field have led to a requirement of the malaria surveillance system to improve the efficiency of delivering the required data. With the technical assistance, the system was reformed to meet these requirements [80]. The modeled malaria information flows before and after the reformation are presented in figure 30. Figure 31 shows the past (before reformation) and present structure of the Cambodian malaria surveillance system. The notable transformation in the system include, but not limited to, the following:

- *Reformation of the national health information system:*

  In response to the increasing demand for more convenient data management and accurate on-time information, the national health information system was reformed to the web-based system from the paper-based using the same data points of the previous system [46]. Although not specifically intended for the malaria program, this reformation provides a significant opportunity for prompt and direct access to facility-level malaria data. Prior to this reformation, there was no reliable flow of data from MoH to CNM, which is a national institution that plays an important role in the national malaria program. As such, it needed to develop its own system for data collection by requesting the HODs for additional reporting tasks [80]. The present web-based system improves the efficiency of data collection and frees them up to concentrate on data-quality issues.

- *Expansion of the VMW program:*

  Since its introduction in the early 2000s, the VMW program has been scaled up in Cambodia [116-117]. VMWs are primarily tasked with reporting data on each rapid diagnostic test including patient information, test results, any treatments, and information on whether the patient were referred to the healthcare providing facilities [42]. Furthermore, they provide basic treatment, conduct active case detection, and track malaria patients [101]. Previously, the individual case data collected by the VMWs were reported in a paper form and directly sent to the national program (NP). However, the flow of data from the VMWs increased exponentially; it was difficult to process the data using the manual data aggregation method [80]. With the development of a new malaria information system (MS), these data are incorporated into the process of the malaria surveillance system.

- *Development of the MS:*
  The MS was developed to process data obtained from the VMWs, health facilities, and bed net distribution status. After the short pilot period, the MS was quickly decentralized and installed at all the targeted HODs. The malaria-case data of the patients of the VMWs and public health facilities are entered into a simple database at HOD level. The data are automatically applied to the national database on a monthly basis [80]. The MS helps in not only improving the efficiency of the data flow, but also in accessing the comprehensive information on the malaria status, which integrates the facility-level malaria data from the web-based national health information system with the VMWs and bed net data obtained from the MS.

Despite having made these efforts and improvements to the system several challenges remain. First, multiple relevant sectors and stakeholders are related to the system. The data collection and validation activities are fragmented with limited coordination both in and outside the health sector [44]. Information from private sectors (PRs) is still limited despite the fact that more than half of the patients seek treatment in the PRs concerning primary health [108, 115]. Second, the quality (accuracy) and reliability of the data collected at health facilities continue to be of concern. One of the possible reasons for this data discrepancy is the inflation of the data at the health facilities with fewer financial resources and supervisory visits. In the validation study on the health information system conducted in 2002, there were no systemic or constituent patterns of data inflation across the country. This suggests that even a small intervention to prevent data inflation at the local facility level could have significant impact on the data quality [44-46]. Ideally, interventions addressing all of these issues at the problematic instances may be effective but not realistic considering the limited healthcare resources. Hence, it is critical to understand the conditions and points of intervention in an on going manner of system operations.

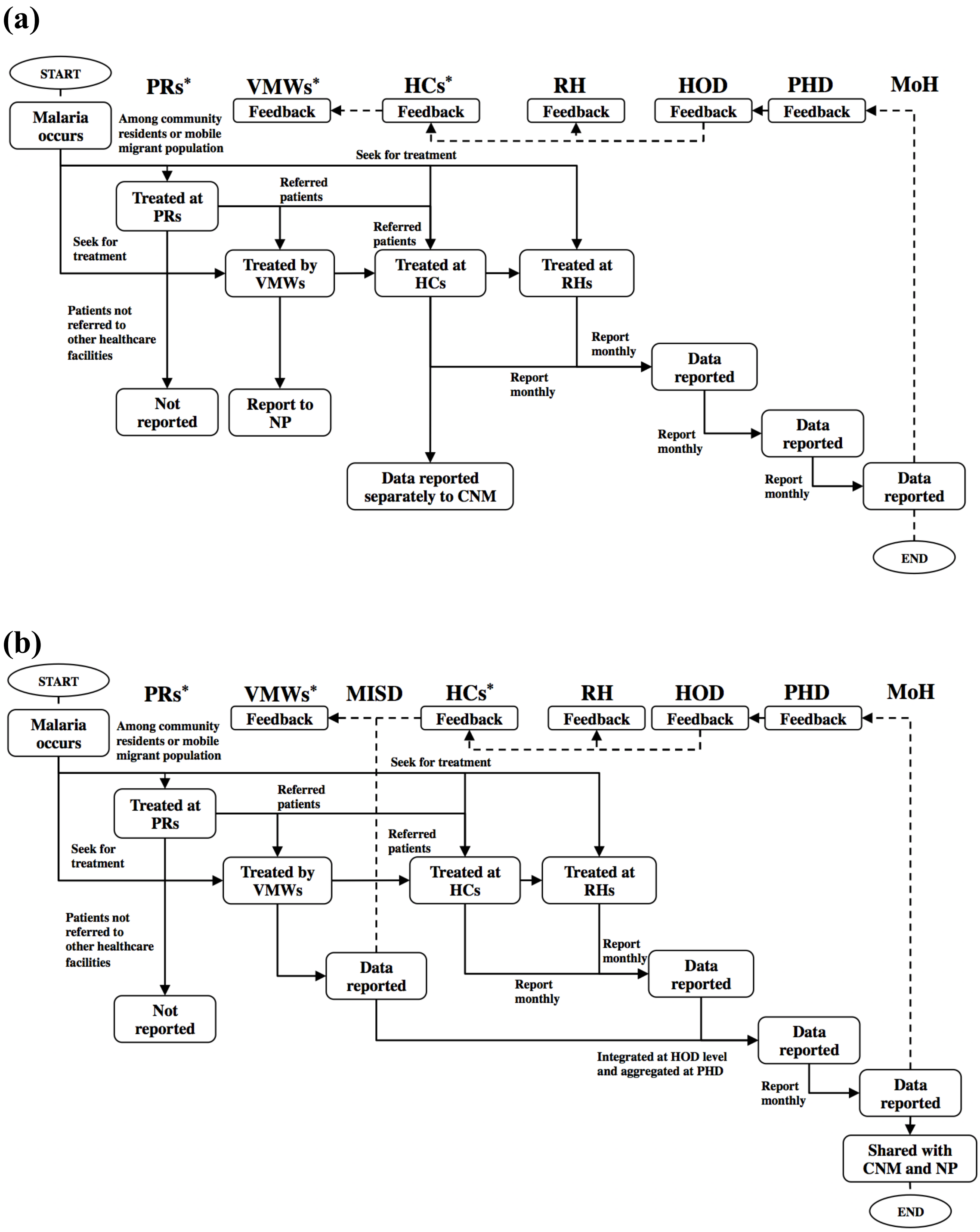



**Figure 30 Modeled malaria information flow before (a) and after (b) the reformation**

MoH: Ministry of Health, HC: Health Center, RH: Referral Hospital, HOD: Health Operational District, PR: Private Sector, PHD: Provincial Health Department, MISD: Malaria Information System Database, CNM: National Center for Parasitology, Entomology and Malaria Control, NP: National Program

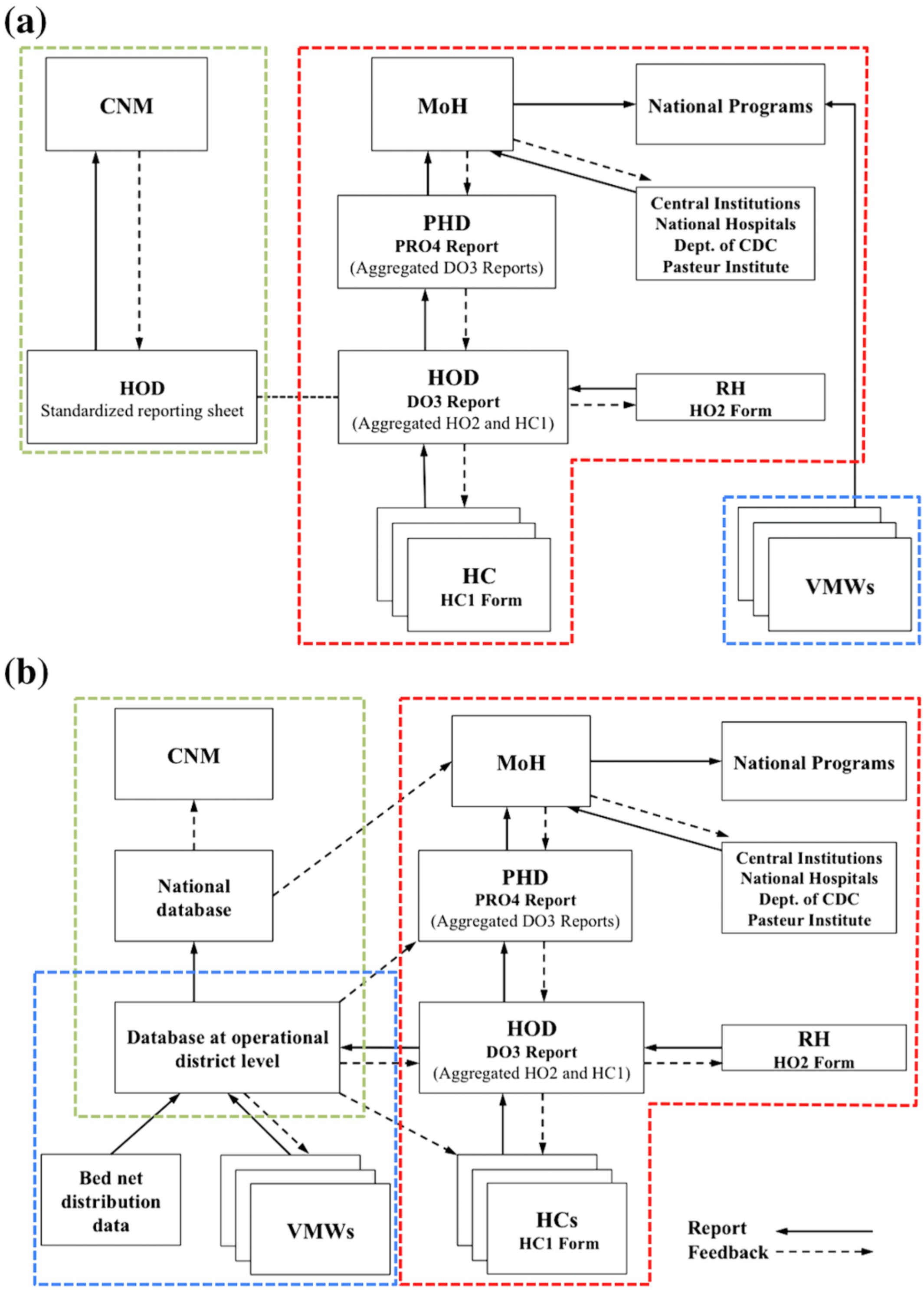



**Figure 31 Past (a) and present (b) structure of malaria surveillance system**

Major differences of the structure between each time interval are highlighted by the dotted rectangle. The red-colored rectangle surrounds the national health information system, the green-colored rectangle presents the reporting pathway from HOD to the national institution, and the blue-colored rectangle shows the changes of case reporting from VMWs because of the introduction of MS. MoH: Ministry of Health, HC: Health Center, RH: Referral Hospital, HOD: Health Operational District, PHD: Provincial Health Department, CNM: National Center for Parasitology, Entomology and Malaria Control, VMW: Village Malaria Worker, MS: Malaria Information System.

### 4.2.3 Malaria-surveillance SoS

Before proceeding to the system modeling and analysis, whether the malaria surveillance system is actually a SoS was examined. The system was characterized based on the five properties (autonomy, belonging, connectivity, diversity, and emergence), proposed by Boardman and Sauser [118-119]. The system under investigation is a system that aims to provide updated and precise information required for evidence-based decision making and day-to-day clinical management. Considering the ever-changing endemic situations, no single system can cover the data requirements. Multiple recording and reporting systems are integrated into a system to prevent duplication for reducing workload of staff. The reliability of the data is enhanced by using standardization and simplification. With the reformation of the national health information system from paper-based to web-based system and the development of the MS, the facility-based data and the report from VMWs are integrated to provide comprehensive data for the purpose of mitigating malaria spread (belonging). Despite the geographically isolated constituent systems, they are interconnected by data and resource flow (connectivity). They are engaged in different target and objectives fro those of the vertical health information system, but brought together for the purpose of mitigating malaria spread (autonomy). Various stakeholders are involved in the system, which allows for the diversity and flexibility in capability of the system (diversity). Considering that many of malaria patients consult with PRs such as community pharmacies and clinics or VMWs as a first medical contact and do not take further treatments at healthcare facilities, the surveillance system can be strengthened with the involvement of PRs and VMWs by facilitating appropriate referral of patients to the healthcare facilities and reporting to the system.

## 4.2 Proposed approach for modeling and analysis

### 4.2.1 Engineering Systems Multiple-Domain Matrix

Engineering Systems (ES) is an interdisciplinary field of study that seeks solutions for important, large, complex sociotechnical problems [120]. ES is characterized by highly complex entanglement of technical and social elements [121] and is conceptualized as open systems that interact with environment. Both the components and environment of ES can change over time. Not only the components themselves but also their properties can change, be added, or dropped. These emergent properties of the systems are characterized by the interactions between the social and technical components [114]. ES-MDM was proposed as an organizing modeling framework, which aids in the conceptualization of ES [122]. In addition, ES-MDM represents a system as it changes over time. The methodology of the ES-MDM is an extension of the design structure matrix (DSM) [123-124] and the domain mapping matrix (DMM) [125] methods, which contain multiple domains and relationships within the system at any given point of time. There are six domains (environmental or system drivers, social or stakeholders, functional including objectives and functions, physical or objects, process or activities, and temporal) that are important to describe ES using ES-MDM. This information can be organized and used to facilitate network and graph-theory analysis [126] for quantitative analysis. Based on the concept of the epidemiological triad [77] of malaria transmission (see Figure 11), the environment surrounding human communities is closely related to the disease causation and should be carefully considered while optimizing system management. As such, the ES-MDM was applied to system modeling and analysis to carefully examine the interactions of multiple domains. In this study, the focus was mainly placed on the interaction of process, architecture, and environment (or system drivers) based on the concept of epidemiological triad. However, the model can include several other viewpoints originated from the interaction of other domains.

### 4.2.2 Modeling malaria surveillance system in Cambodia

In this section, the modeling approach is presented for the malaria surveillance system in Pailin province, which is located in the northwestern region of Cambodia and is one of the places where the issue of drug-resistant malaria was reported [32]. Within the ES-MDM modeling framework, the DSM of the process and architecture domains were examined using the DMM at the intersection of the DSMs. The key deliverables of the system under investigation are quality (particularly for accuracy) and timeliness of the malaria surveillance data. Hence, to optimize these deliverables, the focus was on the flow of the reported data for the modeling. The DSMs for each domain were built using the information collected through a survey of published literature [101], official public documents (guideline and presentations) [45, 80], and interviews of stakeholders such as staffs at regional HCs. Figure 32 and 33 show the multiple-domain matrix of the process and architecture of the system before and after the reformation, respectively.

The abbreviations used in the figures are as follows:

| | |
|---|---|
| CNM | National Center for Parasitology, Entomology and Malaria Control; |
| MoH | Ministry of Health; |
| NP | National Program; |
| DPHI | Department of Planning and Health Information at MoH; |
| MS | Malaria Information System; |
| PHD | Provincial Health Department; |
| HOD | Health Operational District; |
| RH | Referral Hospital; |
| HC | Health Center (HC-1, …, HC-6); |
| VMW | Village Malaria Worker; |
| PR | Private sector, e.g., community pharmacy, private clinic and hospital; |
| MMP | Mobile Migrant Population; |
| HIS | National Health Information System. |

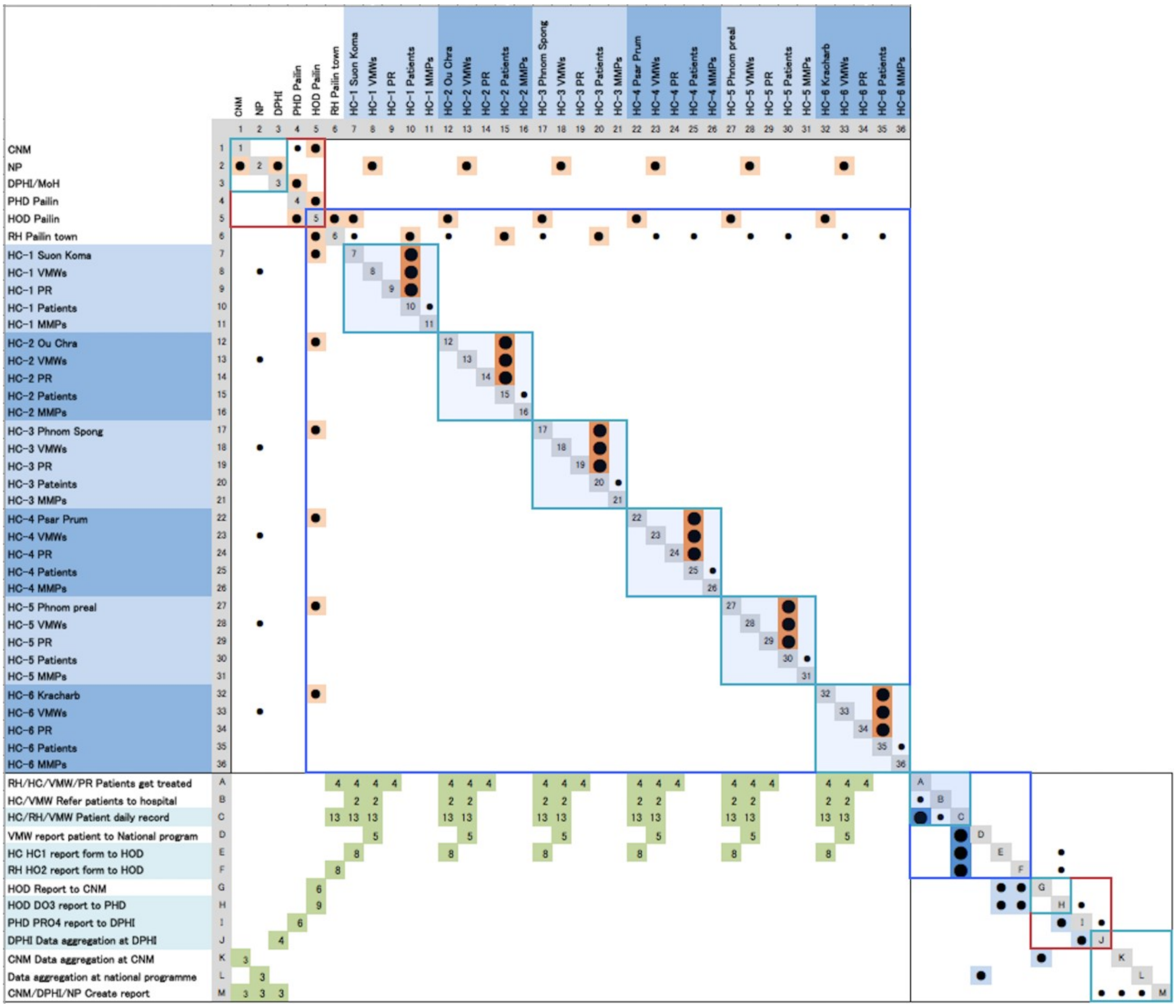


**Figure 32 Multi-domain matrix (before the reformation) of the routine malaria surveillance system in Pailin, Cambodia**

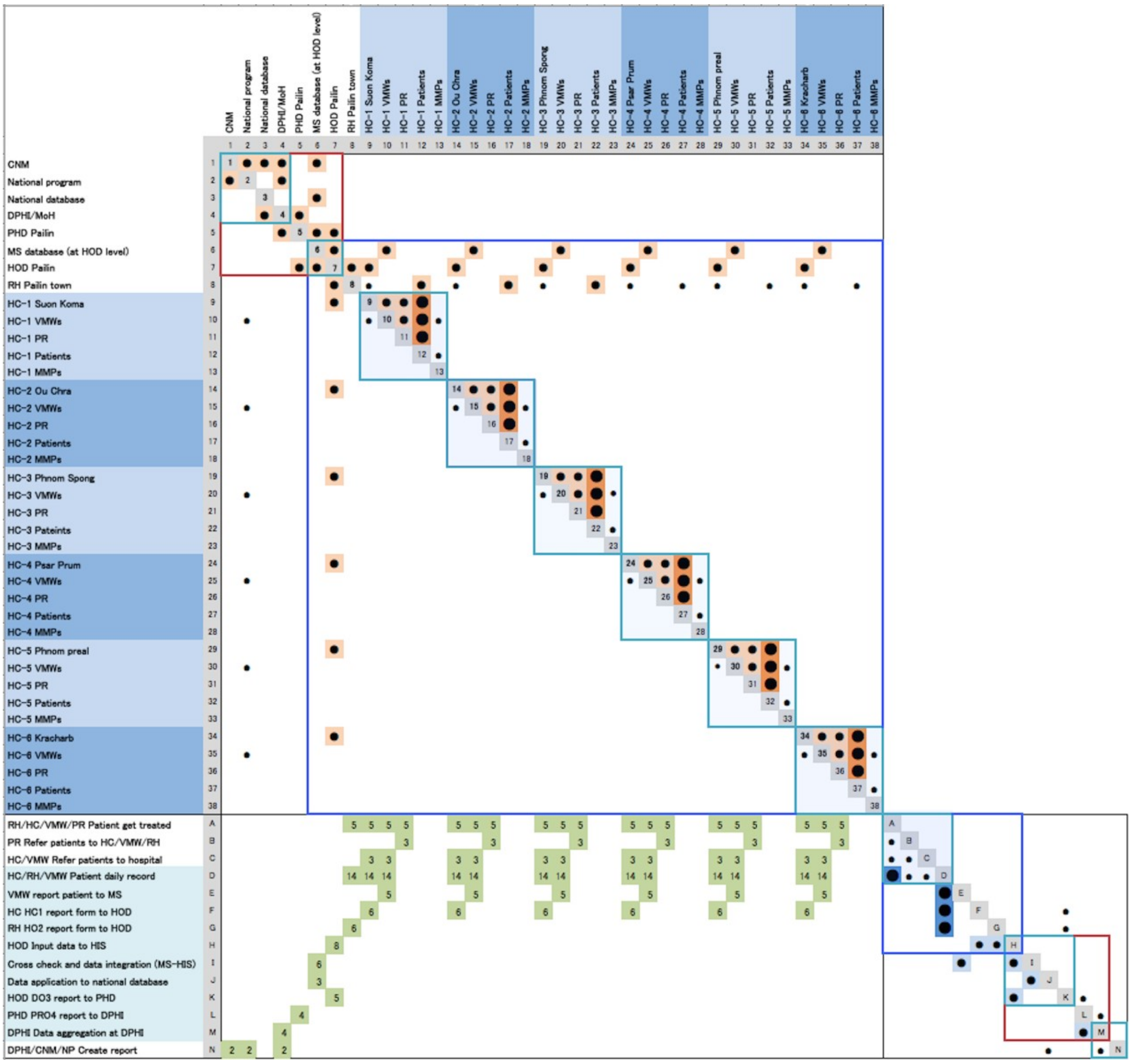


**Figure 33 Multi-domain matrix (after the reformation) of the routine malaria surveillance system in Pailin, Cambodia**

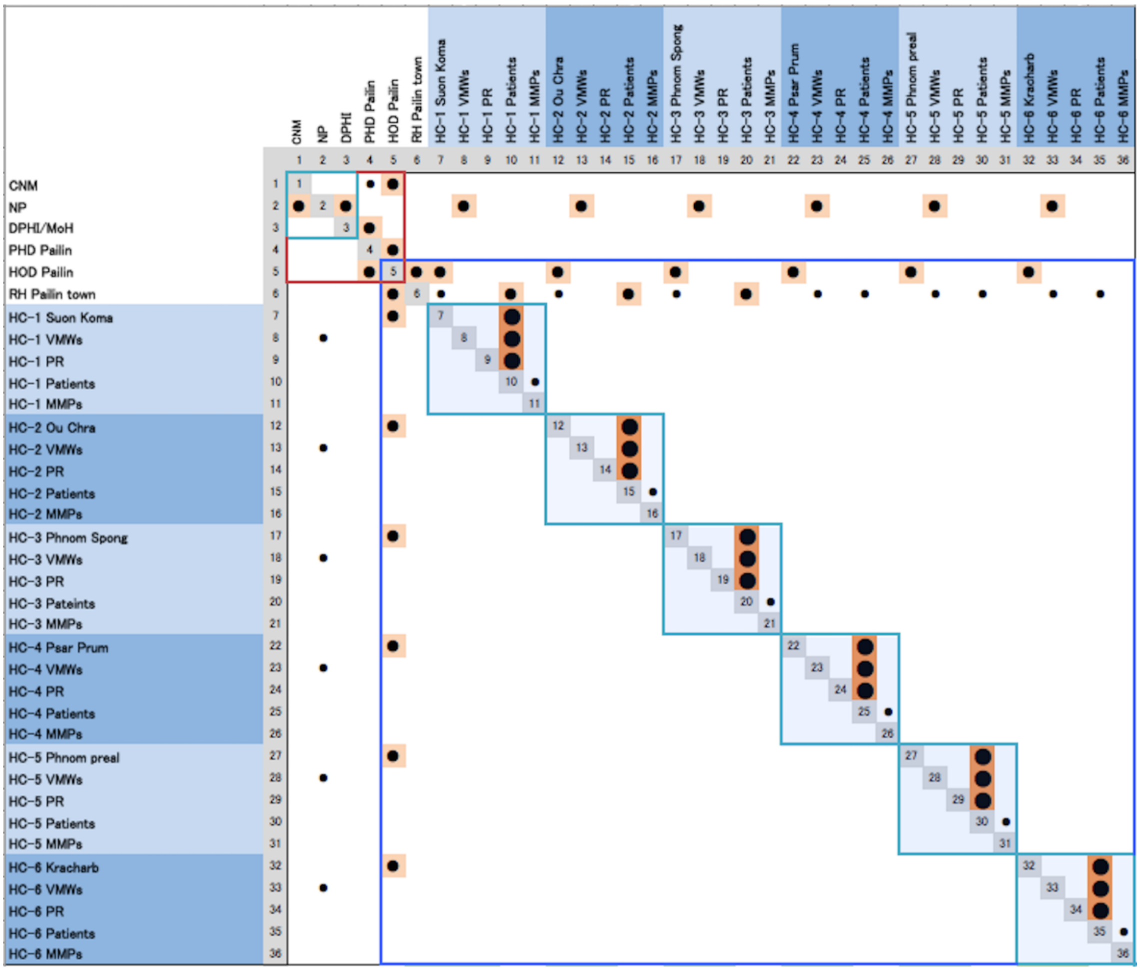


**Figure 34 Architecture domain DSM (before the reformation) of the routine malaria surveillance system in Pailin, Cambodia**

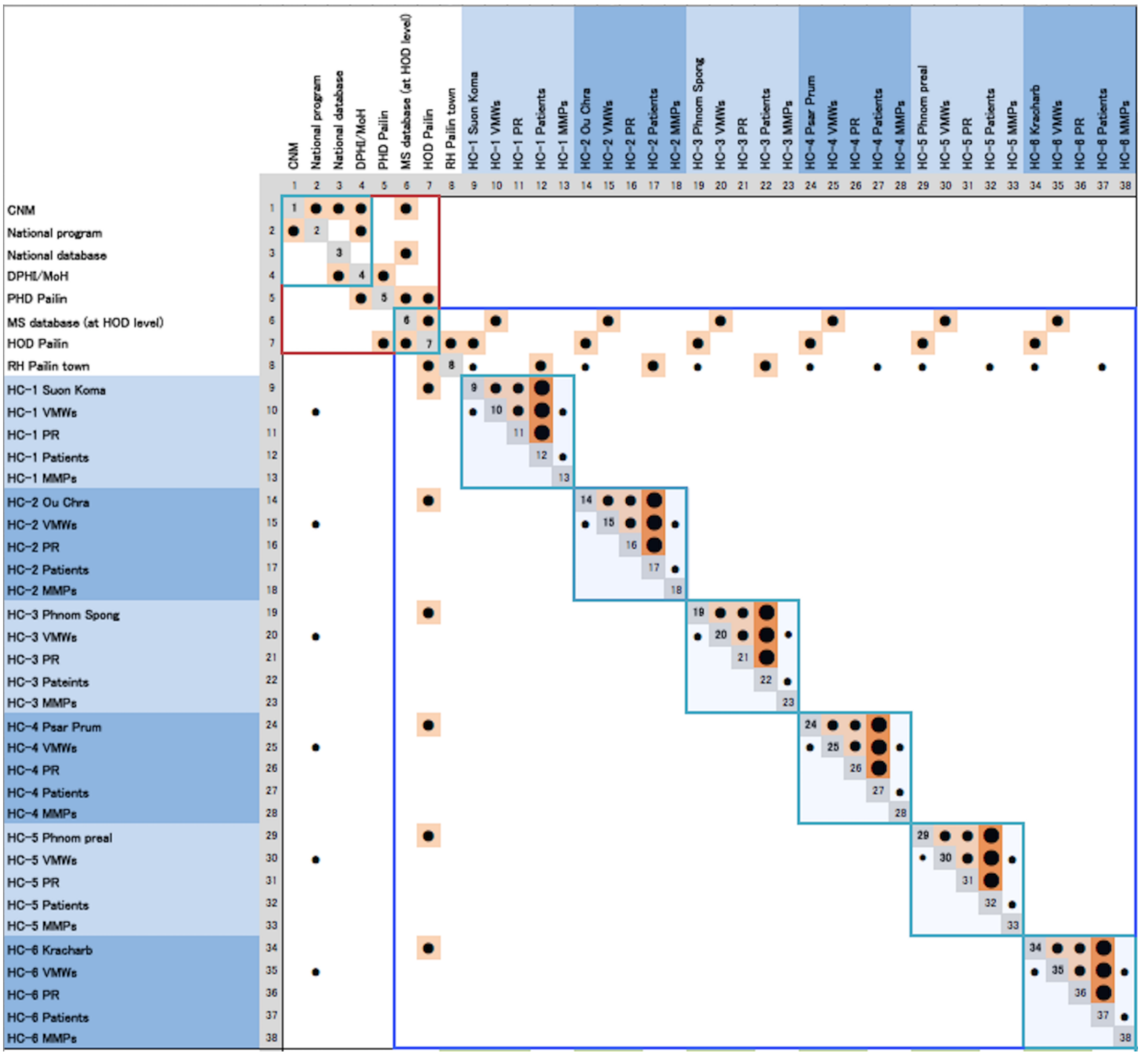


**Figure 35 Architecture domain DSM (after the reformation) of the routine malaria surveillance system in Pailin, Cambodia**

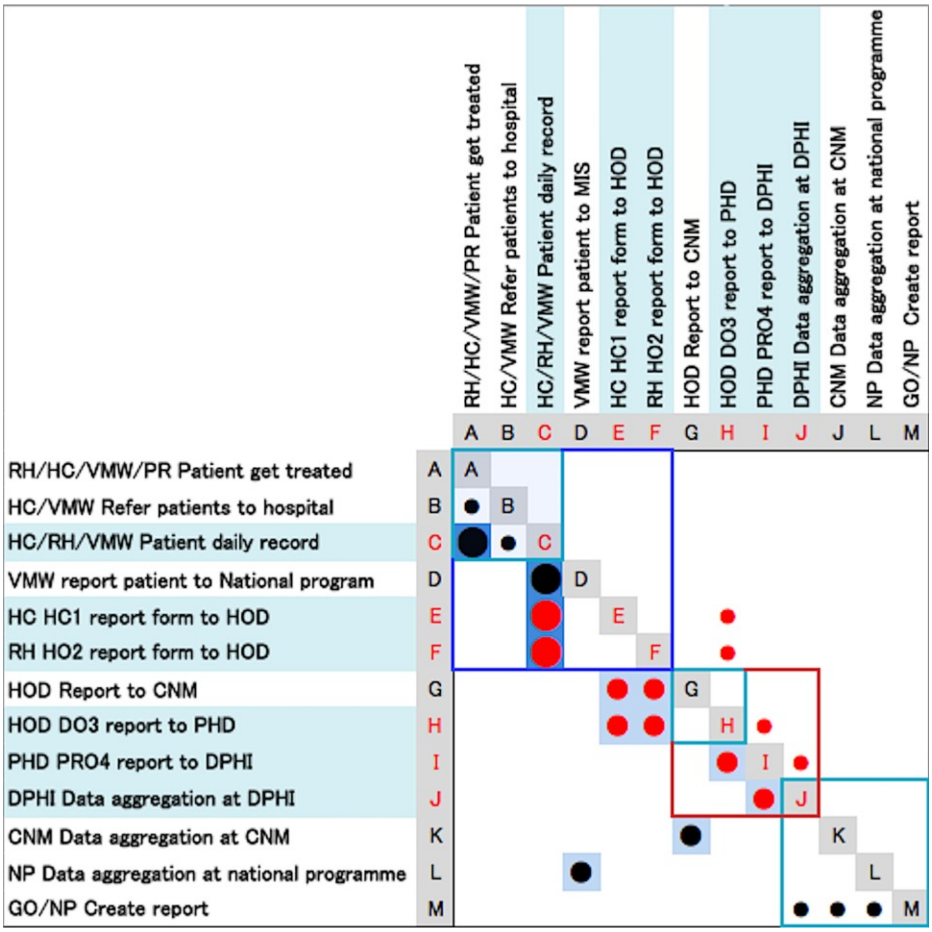



**Figure 36 Process domain DSM (before the reformation) of the routine malaria surveillance system in Pailin, Cambodia**

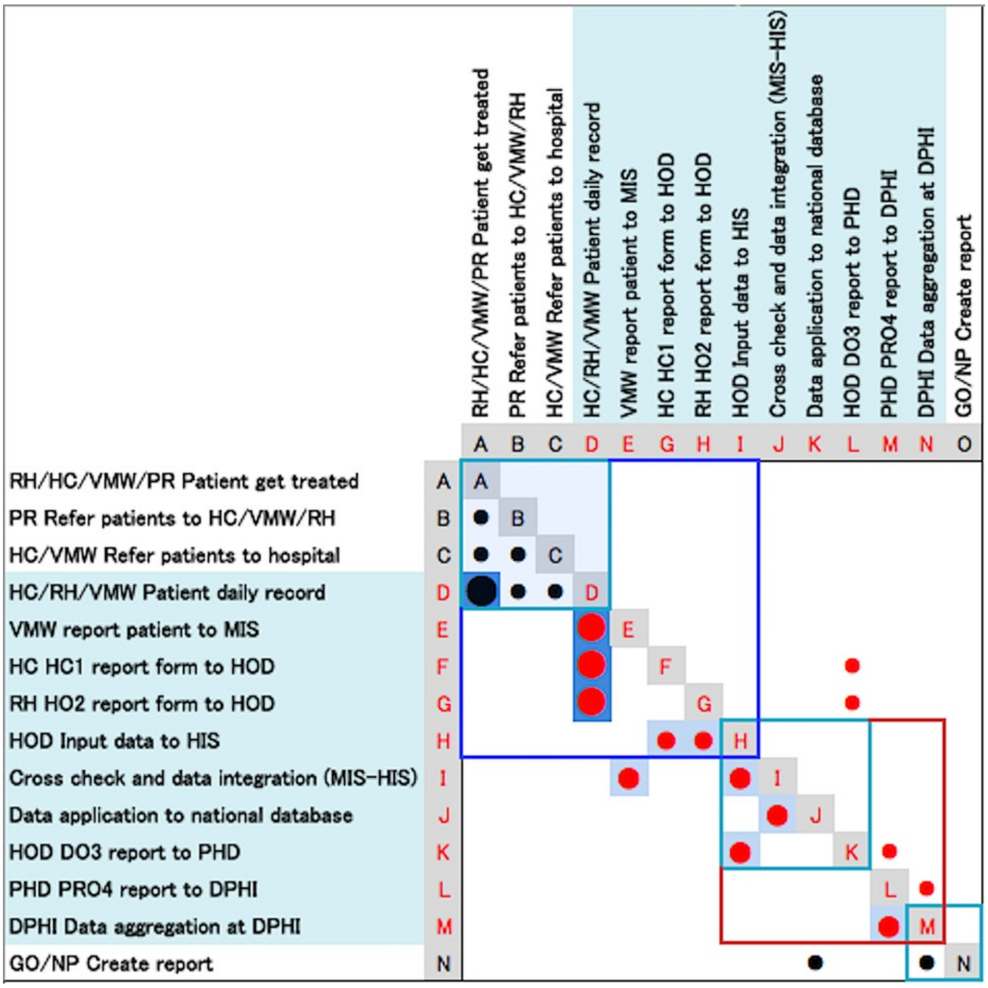



**Figure 37 Process domain DSM (after the reformation) of the routine malaria surveillance system in Pailin, Cambodia**

The DSM in the top left and right present the architecture (Figures 34 and 35) and the process domain (Figures 36 and 37) of the system, respectively. As shown in both the two matrices, connections between the architectural and process components are indicated by circles of different diameters in the off-diagonal cells. They are classified or weighed in accordance with the level of intensity or frequency of the interactions between them. The size of circles corresponds to those levels, i.e., the higher the level of interactions of the architectural components, the classification criteria are as follows:

1) More intensive or closer relationship or communications than those characterized in class 2;
2) Relationship within the level of documented official reporting pathway;
3) Not routine or occasional-basis relationship or communications.

Similarly, the criteria for the process components are as follows:

1) More frequent or intensive process than those categorized in class 2;
2) Frequency or intensity at the level of documented official reporting pathway;
3) Not routine or less intensive process.

The blue-colored square frame separates the architectural and process components below and above the HOD level; the red-colored square frame indicates those related to the data reporting paths at the national level. The DMM of the architecture and process domain is presented at the intersection of the DSMs. The DMM provide a more visible traceability of the respective process components with the corresponding architectural components. The connections between the architectural and process components are indicated using the green-colored cells.

By observing changes between each time point, increased coverage of the VMWs to various stakeholders can be observed at the level of field clinical practice along with the increase in the VMW network. Furthermore, the reporting pathway of the surveillance data from VMWs is fully integrated in the HIS after developing the MS. Using the DSMs, these changes can be flexibly and promptly reflected in the model along with the transformation of the system itself or that of the environment. In this case, this study was performed under the condition that the system is operating appropriately as documented. However, it is unlikely that the system maintains such as state continuously in most of the real-world conditions. In addition, the modeling approach presented in this study can consider these deviations from the designated state, which provides more sophisticated or practical deliverables. This advantage may be more beneficial for open systems with increased complexity, as they are subject to changes along with the way in which they interact with the environment.

### 4.2.3 Analysis of the relative weights of the constituent systems

Next, the relative weights of the constituent systems are analyzed to investigate the systemic influence, i.e., the magnitude of the effect on the performance of the whole system. The epidemiologic triad suggests that environment should be considered in addition to the architecture or process factors in order to manage the quality and timeliness of the data. Hence, the relative weights of each constituent systems were investigated by multiplying the three factors as in (18), where $S_i$ is the score of the weight, $P_i$ is the process, $A_i$ is the architectural component, and $E_i$ represents the environmental attributes of constituent system $i$.

$$S_i = P_i \cdot A_i \cdot E_i \qquad (18)$$

The process attribute is calculated by adding the inputs and outputs of each process component, which is identified in the process domain DSM. They are considered the task burden derived from the processes. The numbers presented in each green-colored cell in the DMM indicates the attributes of each process component. The multiple process components are linked to one constituent system. Hence, they are calculated by aggregating $P_{ij}$, which are attributes of the process component $j$ linked to the constituent systems $i$.

$$P_i = \sum_j P_{ij} \qquad (19)$$

Based on the network of each constituent system and graph theory analysis, network centrality metrics were employed as the architectural attributes. Accordingly, the closeness [127], eigenvector [102], and betweenness centrality $C_i$ of the constituent systems were calculated as

$$A_i = C_{i(closeness,eigenvector,betweenness)} \qquad (20).$$

The risk of malaria transmission was employed to estimate the environmental attributes of the constituent systems. The standardized morbidity ratio (SMR) [81] of the malaria transmission was calculated for estimating the standardized relative risk to compare the data from differently conditioned components as in (21), where $o_i$ is the observed number of cases and $e_i$ is the expected number of cases at the constituent system $i$.

$$E_i = SMR_i = {}^{o_i}/_{e_i} \qquad (21)$$

In this case, the environmental conditions may be highly variable at the HC or contained further peripheral levels. Given the condition that the risk associated the environmental conditions is equal in the HCs and associated lower levels covered by the respective HCs and associated lower levels covered by the respective HCs, the average values of SMR around each HC were calculated using the risk map of the SMR in Pailin [103]; otherwise, $E_i$ is considered to be 1. For interpretations and comparisons across the scores calculated by the different centricity metrics and each time interval, the calculated scores were normalized to relative weights to the maximum score, i.e., by min-max normalization, within the same centricity metrics and time point, whereby the changes due to the system transformation ($\Delta S_i$) can be measured. Table 7 lists the scores of the relative weights of the key constituents before and after the reformation. We employed the eigenvector centrality as a measure of the influence of each network node to investigate the relative weights, i.e., the systemic influence of the constituent system. The decentralization of the weights from the higher data-aggregating layers to the peripheral field practitioners, such as HCs and VMWs, can be observed after the system reformation. These scores can be visualized in the complex network within the system, which may facilitate the common understanding and shared decision making among multiple stakeholders. Figure 38 shows an example of the visualized network using the scores calculated by employing the eigenvector centrality values. Visualization of these network models was conducted using R software (https://www.r-project.org). Implications of results by comparing the different centrality measures are discussed in the next section.

**Table 7 Scores of relative weights for key constituents**

| Constituents | *Before | After | ΔS |
|---|---|---|---|
| CNM | 0.065 | 0.012 | -0.053 |
| DPHI/MoH (DPHI) | 0.03 | 0.039 | 0.009 |
| PHD Pailin (PHD) | 0.104 | 0.065 | -0.039 |
| HOD Pailin (HOD) | 0.772 | 0.572 | -0.2 |
| RH Pailin town (RH) | 0.729 | 0.596 | -0.133 |
| HC-1 Suon Koma (HC-1) | 0.319 | 0.284 | -0.035 |
| HC-1 VMWs (VM1) | 0.068 | 0.148 | 0.08 |
| HC-2 Ou Chra (HC2) | 0.302 | 0.269 | -0.033 |
| HC-2 VMWs (VM2) | 0.064 | 0.14 | 0.076 |
| HC-3 Phom Spong (HC3) | 0.546 | 0.485 | -0.061 |
| HC-3 VMWs (VM3) | 0.116 | 0.253 | 0.137 |
| HC-4 Psar Prum (HC4) | 0.701 | 0.701 | 0 |
| HC-4 VMWs (VM4) | 0.229 | 0.356 | 0.127 |
| HC-5 Phnom Preal (HC5) | 1 | 1 | 0 |
| HC-5 VMWs (VM5) | 0.327 | 0.507 | 0.18 |
| HC-6 Kracharb (HC6) | 0.644 | 0.644 | 0 |
| HC-6 VMWs (VM6) | 0.211 | 0.327 | 0.116 |

*Architectural attributes are calculated using the eigenvector centrality value

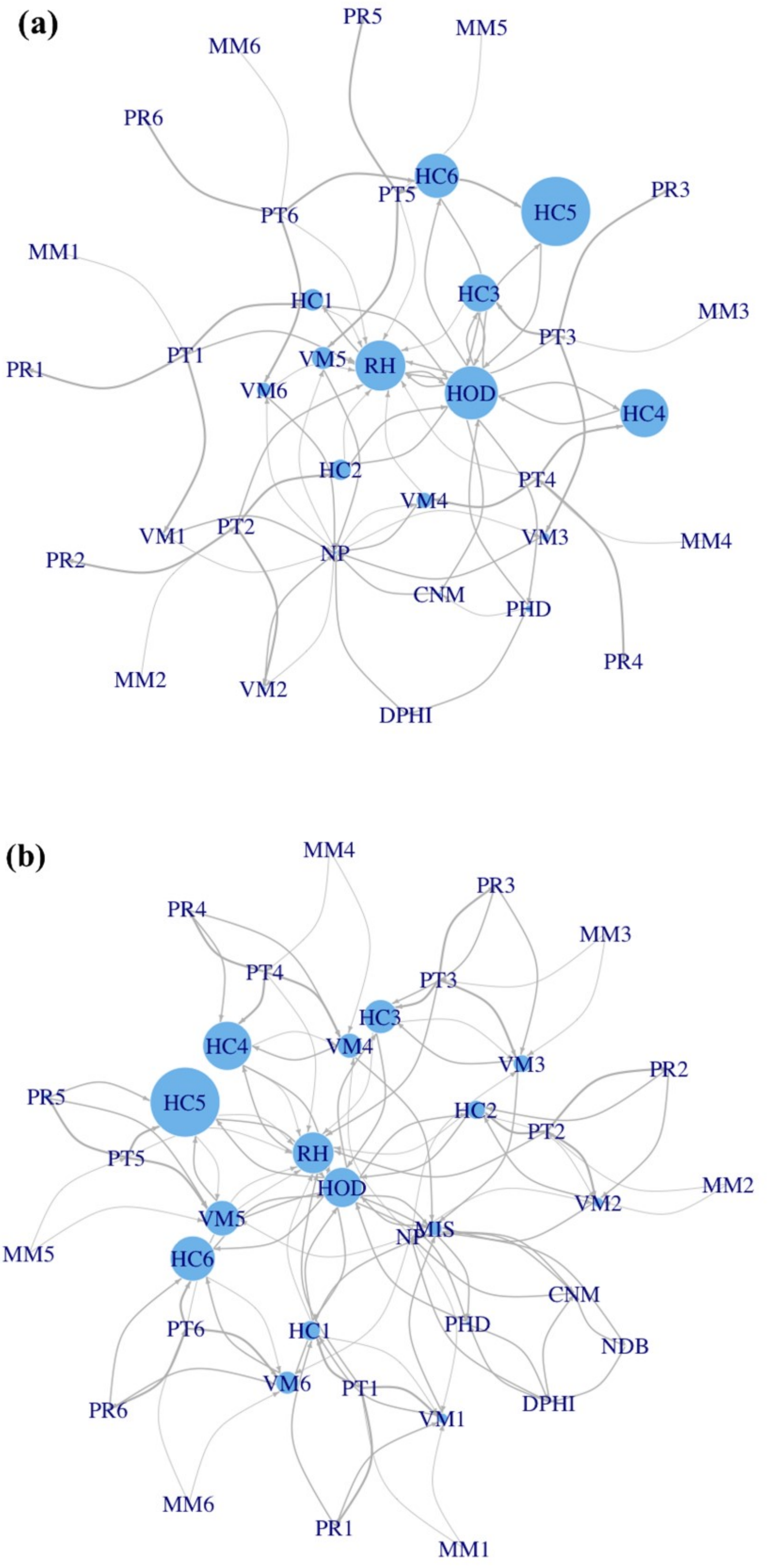



**Figure 38 Example of the visualized network using scores calculated by employing the eigenvector centrality values (a) before and (b) after the system reformation**

## 4.3 Validation

### 4.3.1 Capturing the transformation of the SoS

As a part of validation of the proposed approach, the dataset of the absolute values of $\Delta S$ ($|\Delta S|$) was resampled 2000 times to create confidence intervals by employing the Corrected non-parametric bias and accelerated percentile method [99] to assess the distribution of $|\Delta S|$. Figure 39 shows the kernel density plots of the resampled $|\Delta S|$. The 95% confidence intervals are 0.007 – 0.037, 0.029 – 0.063, and 0.048 – 0.142 for the scores calculated by the closeness, for the eigenvector, and for the betweenness centrality, respectively. The plots do not include zero in the 95% intervals, i.e., the differences are statistically significant when the null hypothesis $H_0$ $|\Delta S| = 0$ is discarded at a significance level $\alpha = 0.05$. Hence, this approach is assumed to help capture the transformation of the system under investigation at least to a particular extent.

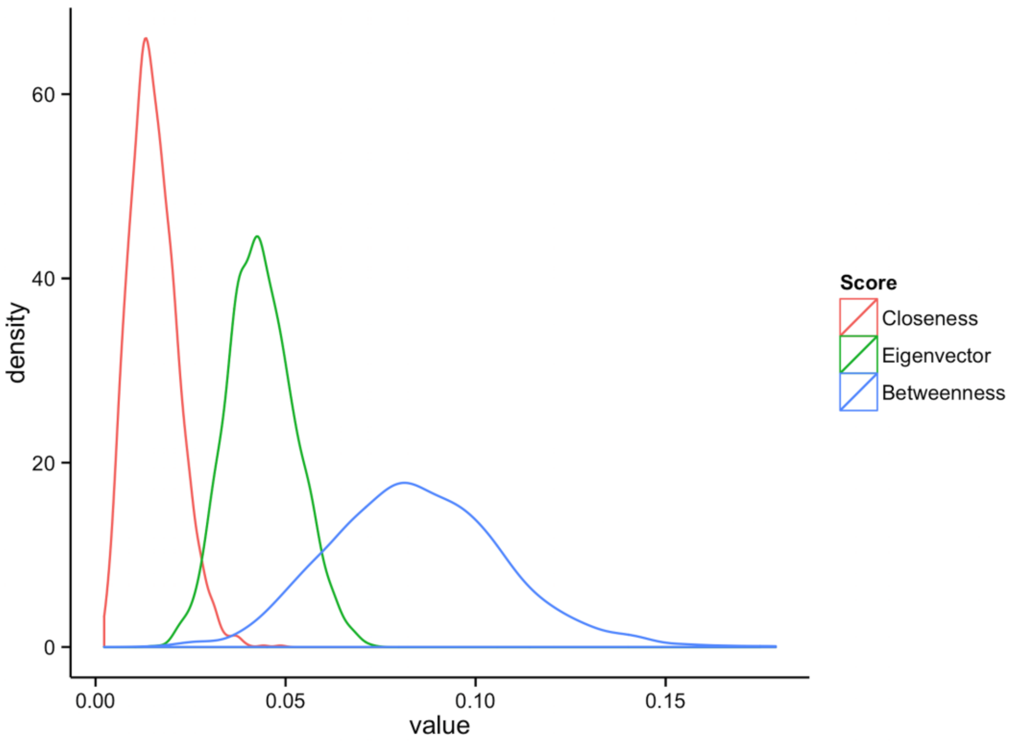



**Figure 39 Kernel density plot of the resampled $|\Delta S|$ before and after the reformation calculated by employing closeness, eigenvector, and betweenness centrality**

### 4.3.2 Simulation test

Next, we discuss whether it can help identify the correct targets of intervention in the health information SoS under investigation. Since, it is difficult to test large and complex SoS in real-world setting, we designed simulation tests to estimate the systemic influence of each constituent system and compared them with scores calculated by the method discussed in the previous section. The SoS and its constituents have several distinctive characteristics such as autonomy and emergence; hence, we employed the agent-based modeling (ABM) simulation [128] to build the simulation environment. ABM is a computational tool to model dynamical systems, by programming each element, and it has a number of applications in several field of study [129]. With the ABM simulation, a system is modeled as a collection of autonomous decision-making entities called agents. Each agent presents its behavior separately by assessing its situation and making decision on the basis of programmed set of rules [128]. Due to its advantages in modeling dynamic system functionality as a collection of autonomous entities, ABM was employed to simulate SoS [119]. In this study, we used NetLogo 5.3.1 [130] as an ABM tool. For comparison of the data obtained, we modeled the flow of surveillance data in the same case of the system post reformation as modeled in the previous section. Within the modeling environment, each SoS constituent was programmed to act for surveillance reporting. Figure 40 shows the flowchart of the processes in the system under investigation. A case of malaria first occurs in the villages covered by each of six HC. Assuming that the case count data can follow a Poisson distribution, the number of patients was calculated as in (22), where, $N_k$ is the actual number of cases in area $k$ (k = 0, 1, …, 6). In this case, $e_k$ was calculated by apportioning the preset total number of patients $N_{total}$ based on the magnitude of environmental attributes $E_k$ of area $k$ (23), which was calculated in the previous section. Likewise, a few malaria patients were found from MMPs, which was calculated with a constant value of $e_k$ by (22) and added to the number of $N_k$.

$$N_i | e_k \sim Po(e_{\ k}) \qquad (22)$$

$$e_k = N_{\ total} \cdot {}^{E_k} / {}_{\sum E_k} \qquad (23)$$

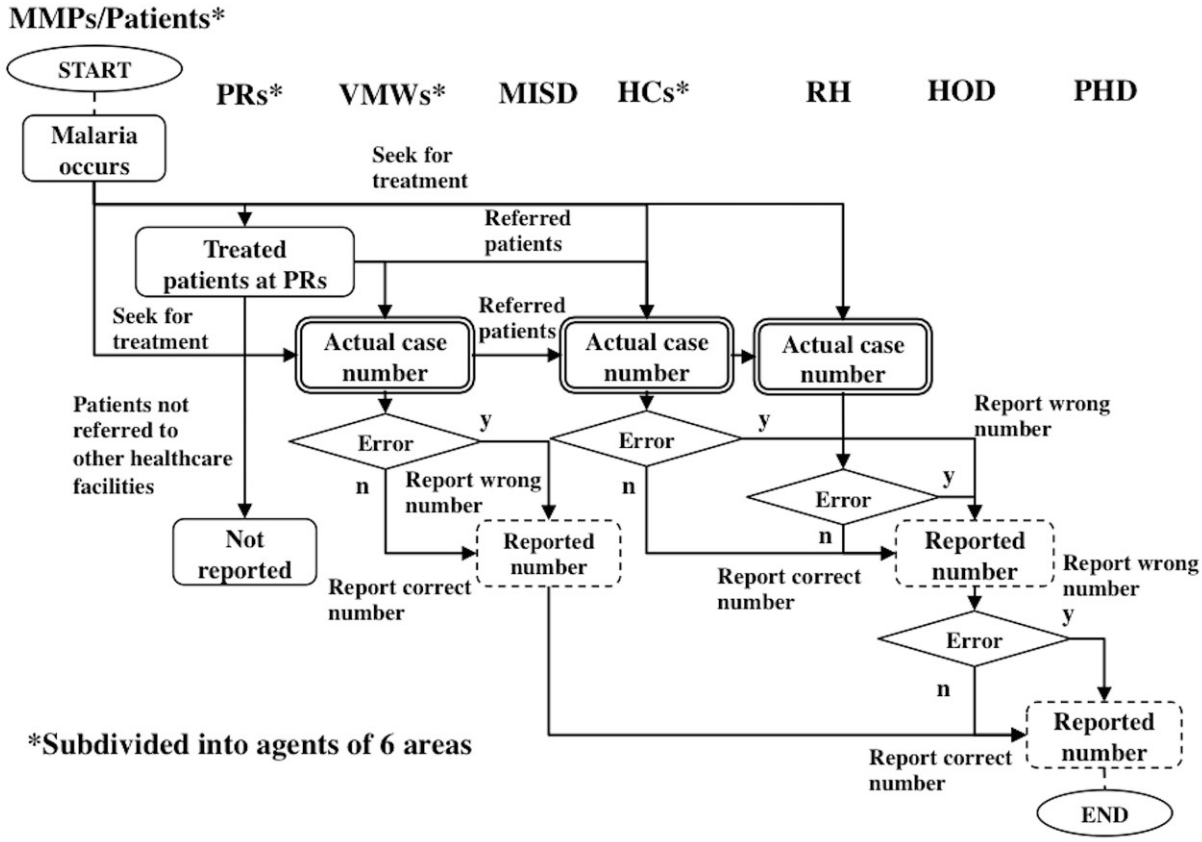



**Figure 40 Flowchart of simulated processes in the system under investigation**

HC: Health Center, RH: Referral Hospital, HOD: Health Operational District, PR: Private Sector, PHD: Provincial Health Department, MISD: Malaria Information System Database, MMP: Mobile Migrant Population

The malaria patients seek treatment at PRs, VMWs, HCs, or RHs. Certain numbers of patients visited PRs as the first medical contacts are referred to HCs or VMWs. In the same way, a few patients treated by VMWs are referred to HCs or RH, and patients in severe conditions among those treated at HCs are referred to RH. Thus, the number of patients $N_k$ was distributed at a certain constant rates to a healthcare facility or a VMW $i$ in the same area. HCs and RHs periodically report the number of treated patients to the HOD, which report the aggregated numbers from the HCs and the RH to the PHD. The VMWs report the number of patients to the MS. However, deviations between reported data and actual treated numbers occur at this stage for several reasons, such as unintentional data collection errors or data inflation. Since the occurrence of errors is unpredictable, we employed the probabilistic method in the model described in [70].

The occurrence and magnitude of deviations $D_i$ were modeled by introducing an indicator variable $R_i \in \{0, 1\}$ specifying whether an error occurred at the reporting process by each reporting agent (24). If $R_i = 0$, no errors had occurred and the correct number was reported, and $D_i = 1$. However, if $R_i = 1$, errors occurred in the reporting process, and a number was selected from the fixed error distribution $\theta_i$. In this study, $\theta_i$ was a uniform distribution over a preset range of the magnitude of deviation.

$$\begin{cases} D_i | R_i = 1, & if\ R_i = 0 \\ D_i | R_i = \theta_i \sim Continuous(\theta_i), & otherwise \end{cases} \quad (24)$$

$R_i$ was conditioned by the probability of the occurrence of errors $\lambda$.

$$R_i | \lambda \sim Bernoulli(\lambda) \quad (25)$$

$\lambda$ is an unknown random variable. Therefore, we defined it as a beta distribution, which is a continuous distribution over the real number from zero to one, for the purpose of modeling its flexibility.

$$\lambda \sim Beta(\alpha, \beta) \quad (26)$$

For mathematical convenience, a combination of a beta prior distribution for a Bernoulli random variable is standard practice in probabilistic modeling [131]. At the implementation of the model, we set fixed constants of two hyper parameters $\alpha = 5$ and $\beta = 1$ of the beta distribution given that error occurrence was not frequent. By multiplying $D_i$ with the actual number of cases $n_i$, the reported number $r_i$ of each agent was calculated (27). Since patients were not treated at the HOD, we considered the $n_i$ of the HOD as the aggregation of reported values $r_i$ from the HCs and the RH, by which the influence of the error at the HOD was calculated while distinguishing the effect of errors caused by the HCs and the RH.

$$r_i = n_i \cdot D_i \quad (27)$$

To evaluate the systemic influence of constituent systems, we measured the correctness of the information reported by each agent. Although the correctness of the information is often regarded as a quality metric, a number of studies have quantified information correctness [74], [132]. In this study, correctness was measured by a function of distance between the actual number of cases $n_i$ and the corresponding reported number $r_i$ as in (28), where $c_i$ is the correctness of information reported by healthcare facility or VMW $i$:

$$C_i = 1 - {|n_i - r_i|}/{n_i} = 1 - |1 - D_i| \qquad (28)$$

Therefore, correctness was dependent on the magnitude of deviance of each reporting agent. We then calculated the systemic influence $S_i$ as in (29) as a function of $D_i$ and the fraction of patients reported by healthcare facility or VMW $i$ in the system:

$$S_i = |1 - D_i| \cdot {n_i}/{\sum n_i} \qquad (29)$$

We set the parameters for the simulation test as shown in table 8, based on the information from a previous report [115] and an interview. Considering $N_{total}$ = 1200 and $e_k$ for MMP at 5 as a constant value, we ran the simulation test for 1000 times and sampled the data from each execution. Since the magnitude of deviation in the reported information consisted of unknown variables sampled from a preset range of $D_i$ to run the simulation as a sensitivity analysis. The different ranges of $D_i$ were set for HCs/RH/VMWs and the HOD considering their different roles in the reporting process. The calculated systemic influence values were normalized to relative weights to the maximum value, whereby the value could be compared with the scores calculated in the previous section.

**Table 8 Preset parameters for the simulation test**

| Parameters | Surrounding area of RH ($k$ = 1–3) | Distal area from RH ($k$ = 4–6) |
|---|---|---|
| Percentage of patients visit PRs* | 50% | 50% |
| Percentage of patients visit VMWs* | 30% | 30% |
| Percentage of patients visit HCs* | 16% | 19% |
| Percentage of patients directly visit RH* | 4% | 1% |
| Referral from PR to VMW | 25% | 25% |
| Referral from PR to HC | 50% | 50% |
| Referral from VMW to HC/HC to RH | 2% | 2% |

*For the first medical contact when seeking treatment for malaria.

### 4.3.3 Results and discussion

In the repeated simulation tests, the actual number of malaria cases treated at HCs and RH was 571.996 ± 15.882, and the number of those treated by VMWs was 528.669 ± 14.71 (mean ± SD). These numbers were in close agreement with figures from a previous report [78]. Figure 41 shows the correctness values of the information provided by the entire system under investigation (global correctness) and the minimum correctness values of the information provided by constituent systems, under the condition that the possible ranges of deviation occurred were -20 to +50% at HCs/RH/VMWs and ±5% at the HOD. The cumulative mean of global correctness and the minimum correctness values were 0.958 ± 0.034 and 0.677 ± 0.137, respectively (mean ± SD). These results indicate the possibility of hidden imbalances in the performance of constituent systems due to the cancelation of the different directions of generated values, even though the performance of the entire system seemed acceptable. This test case presents the importance of the appropriate evaluation of the constituent systems, not only to support decision-making, but also to avoid overlooking significant deviations.

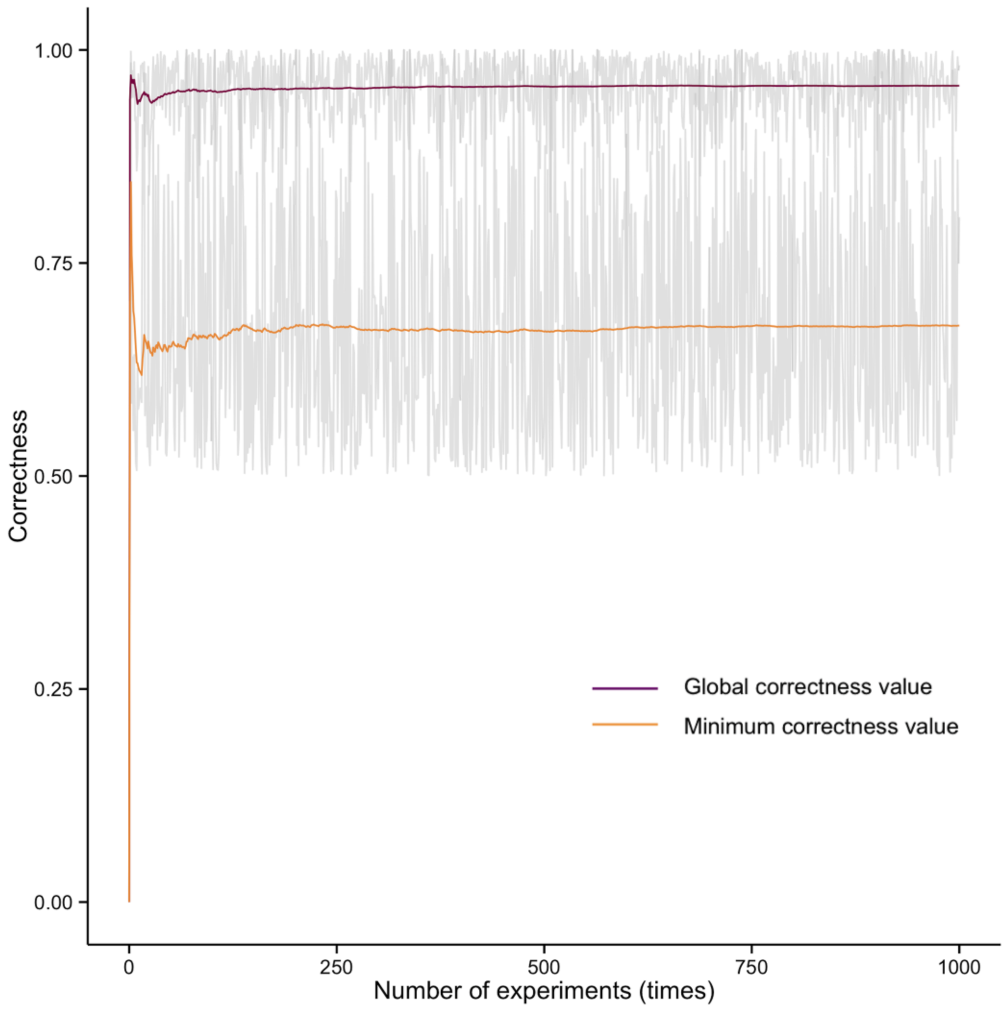



**Figure 41 Calculated global correctness values and minimum correctness values**

Simulation was run under conditions where the possible ranges of deviation were -20 to +50% at HCs/RH/VMWs and ±5% at the HOD. Solid lines show the cumulative mean values.

**Table 9 Scores of relative weights for key constituents**

| Constituents | Systemic influence | Score (eigenvector) | Score (betweenness) | Score (closeness) |
|---|---|---|---|---|
| HOD Pailin | 0.233 | 0.572 | 1 | 0.303 |
| RH Pailin town | 0.2 | 0.596 | 0.676 | 0.792 |
| HC1 Suon Koma | 0.281 | 0.284 | 0.165 | 0.278 |
| VM1 VMWs | 0.267 | 0.148 | 0.114 | 0.244 |
| HC2 Ou Chra | 0.226 | 0.269 | 0.156 | 0.263 |
| VM2 VMWs | 0.24 | 0.14 | 0.108 | 0.231 |
| HC3 Phom Spong | 0.464 | 0.485 | 0.282 | 0.475 |
| VM3 VMWs | 0.442 | 0.253 | 0.195 | 0.417 |
| HC4 Psar Prum | 0.776 | 0.701 | 0.39 | 0.701 |
| VM4 VMWs | 0.774 | 0.356 | 0.208 | 0.616 |
| HC5 Phnom Preal | 1 | 1 | 0.557 | 1 |
| VM5 VMWs | 0.991 | 0.507 | 0.296 | 0.879 |
| HC6 Kracharb | 0.7 | 0.644 | 0.359 | 0.644 |
| VM6 VMWs | 0.657 | 0.327 | 0.191 | 0.566 |

Simulation was run under conditions where the possible ranges of deviation were -20 to +50% at HCs/RH/VMWs and ±5% at the HOD. Solid lines show the cumulative mean values.

Table 9 presents the calculated systemic influence of key constituents and scores of relative weights calculated by employing the eigenvector, betweenness, and closeness centralities for architectural attributes. As shown in the scatter plot of the systemic influence values and the scores calculated by employing the eigenvector centrality of key constituent systems (see Figure 42), several differences can be observed among the values, while those of the HCs were in good agreement. The systemic influence values of the VMWs were consistently higher than the calculated scores. Likewise, the systemic influence values of RH and HOD were lower than those of the calculated scores. Contrasting these disagreements provided several indications for understanding the meanings of the scores. Given that similar numbers of malaria cases were reported by the HCs and the VMWs and the same rule was applied to error incidence, it is reasonable to expect that similar systemic influence values were estimated for them. However, the effect or behaviors of the HCs and the VMWs may be different in the system. As shown in the architectural domain DSM in figure 33, the HCs had a direct connection and make a certain level of influence on the VMWs by providing advice and guidance for treating malaria cases [101]. Assuming the pervasive effect of the systemic intervention from the HCs to the VMWs, it is more appropriate to take these factors into account when evaluating the relative weights of the constituent systems. The same explanation can be applied to the relationship between the HOD or the RH and the HCs.

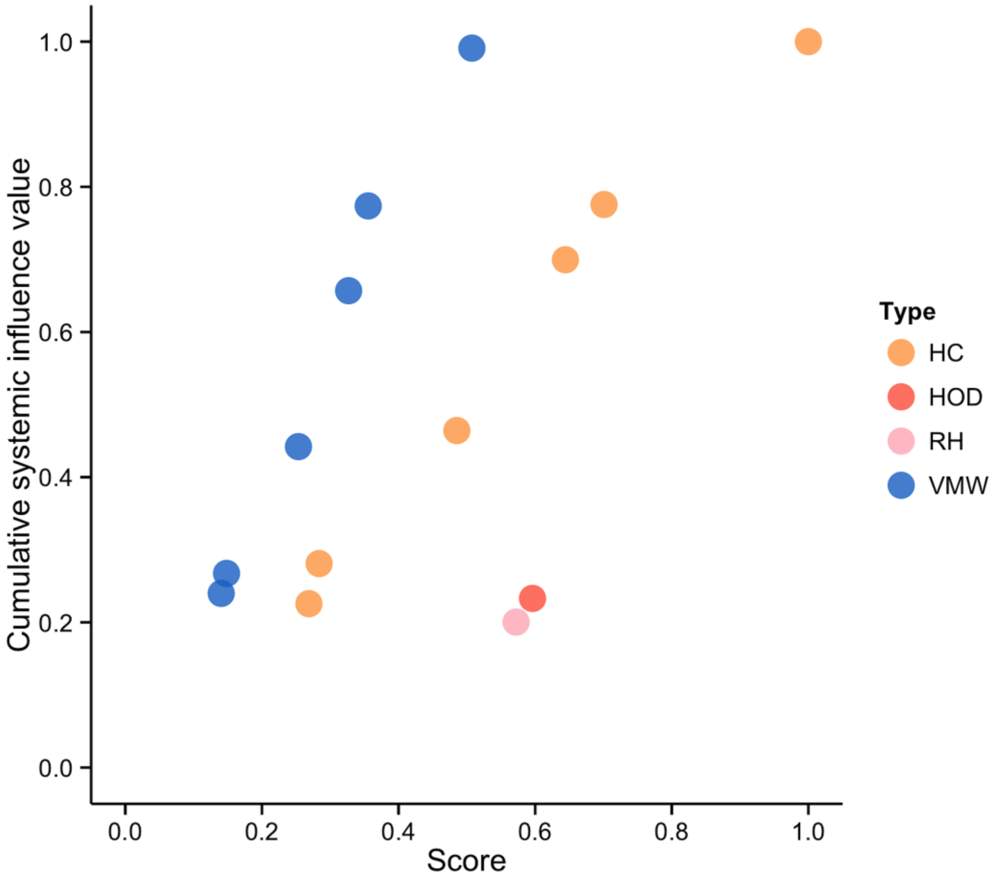



**Figure 42 Scatterplot of cumulative systemic influence value and scores calculated by employing eigenvector centrality of key constituent systems**

Simulation was run under conditions where the possible ranges of deviation were -20 to +50% at HCs/RH/VMWs and ±5% at the HOD. Solid lines show the cumulative mean values.

The architectural attributes $A_i$ as component of the scoring approach can help capture these effects of constituent systems. As demonstrated in the previous work [114], information on the ES-MDM can be organized to facilitate network and graph-theory analysis. This advantage could provide opportunities to obtain more insight by calculating centrality values originated from the architecture of the SoS. However the scores calculated by employing betweenness centrality presented more different patterns from the other two types of values. As betweenness centrality measures the extent to which a network node lies along the shortest paths between other nodes, eigenvector centrality as a measure of influence of each network node can be more appropriately used in the case of measuring systemic influence. The closeness centrality, as a measure of distance between the network node and other nodes, may reflect the influence but the influence of connected other nodes are not considered in this measure. The process attribute $P_i$ also helps identify the important group of constituent systems from the perspective of the process. In this case, the HCs were estimated to be the highest scores when considering both $A_i$ and $P_i$, whereas the intermediate score was estimated to the HOD. The sensitivity analysis of the simulation test suggested consistent findings as indicated by the difference in the slopes of the effect when changing the corresponding values at least within the simulated range (see Figure 43). Once the important group of constituent systems are identified, prioritization within the group can be carried out by the environmental attribute $E_i$, as indicated by the satisfactory agreement between the values of the simulation test and the scoring estimates of the HCs. Hence, the appropriate modeling of the system and the selection of architectural and process measures are critically important and may determine the applicability of this scoring approach to the system in other domains. In the same manner, other factors related to environmental risk, e.g., the status of drug resistance, immunity, or susceptibility of local residents, can be considered when formulating $E_i$ to reflect more precise risk.

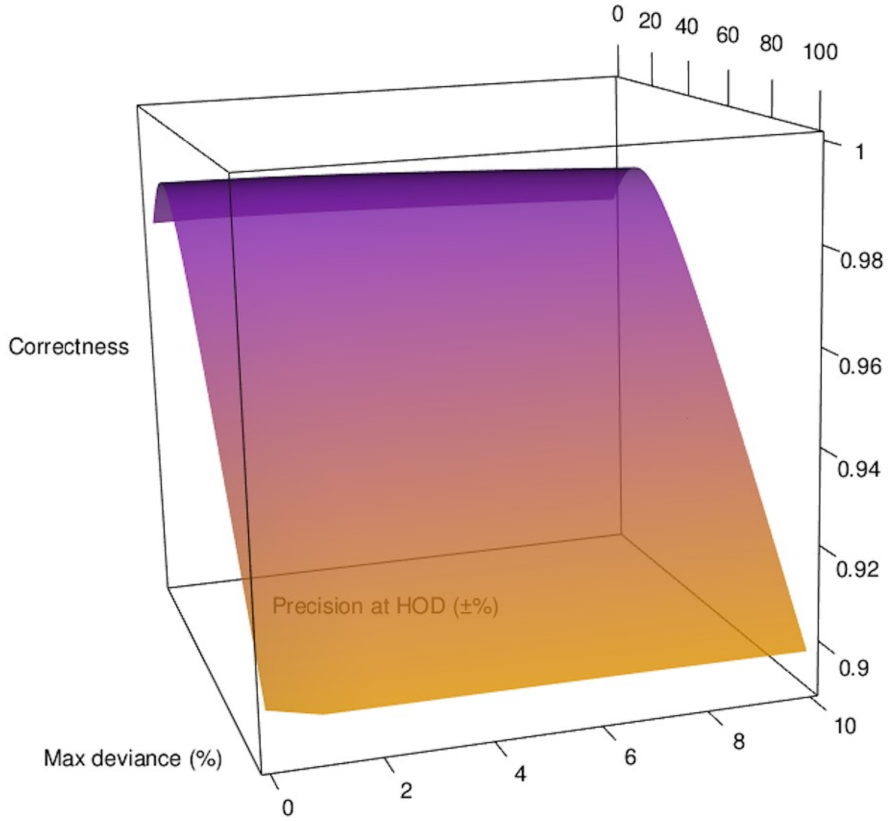



**Figure 43 Sensitivity analysis of the simulation test [global correctness values at any given pairs of max deviation at HCs/RH/VMWs (%) and precision at HOD (±%)]**

Although these results provide further insights of the scoring approach, other test option such as dynamic systems theory should be compared before reaching to more conclusive results. The evolutionary game theory could also be applied and proven effective for this type of problem since they are difficult to be analytically solved [133-134]. Moreover, the model of the system presented in this study is merely a result of a viewpoint of stakeholders. The level of detail and abstraction of the model may differ based on the requirement or viewpoint of the modelers. However the model can readily accommodate changes to the system architecture and process, which is convenient for managing transitional complexities of an open system that transforms over time, such as the ES. The scoring approach proposed here is not suitable for calculating the absolute importance of the constituent systems. However, it is clear that the relative importance of the constituent systems is also significant in several situations, such as planning optimal resource allocations with limited resources.

In the context of this SoS, the modeling and scoring approach presented in this study would highlight to a decision-maker some constituent systems (HC5, HC4, HC6, RH, HOD etc.) as the candidate for improvement efforts as they appeared to have a great deal of influence in the overall effectiveness of the SoS. This SoS transforms over time due to the changes in the various aspects of the system such as the architecture, the process of data collection, and the changes in the environmental risk along with the progress of malaria containment. The modeling and scoring approach can readily accommodate such changes and provide guidance to a decision-maker for the resource management and allocations of improvement efforts on a timely basis. There are a number of applicable patterns of interventions in the large complex health information SoS. Accordingly, by the scoring approach presented here, temporal and continual guidance can be provided to decision-makers, which can lead to an improvement in the quality of data and eventually the level of public health.

## 4.4 Chapter summary

The health information system is a SoS that transforms over time in an attempt to improve its efficiency by considering the effect of the environment. In this study, the process and architecture of a malaria surveillance system were modeled using the ES-MDM modeling framework. The models help examine the interrelationship between the process and architecture of the constituent systems while providing visible insights over multiple domains to various stakeholders. Moreover, the models can consider necessary revisions in an iterative manner, which can help reflect practical situations. Using the attributes of the process, the architecture, and the risk associated with environment, the relative weights of the constituent systems are scored at each interval of time. As indicated in the confidence intervals of the absolute differences in the scores, this approach yields the transformations of the system under investigation to particular degree. By comparing the scores with the results of the simulation test using ABM, further insights into the process, the limitations, and advantages of the scoring approach were discussed. This approach provides the first step in analyzing the transitional conditions of constituent systems as well as the entire health information SoS, whereby the informed decision-making for optimizing the continuous system management is facilitated. With this initial step for the scoring of the relative weights of the constituent systems, the study can be expanded to examining cases of effective use of scores in optimizing healthcare resource allocation. Since the scores capture the transitions of systems while continually providing the guidance to decision-makers, the application of the estimated scores to the dynamic resource allocation problem for more efficient resource distribution is expected. Furthermore, the studies to expand the model to include other domains in the ES-MDM framework such as social and functional domains to obtain further insights from interaction between multiple domains are considered. Through these investigations, further progress can be made in the health information SoS engineering, leading to improvements in providing information fundamental to health.

# 5. TRANSITIONAL COMPLEXITIES OF HELATH

# INFORMATION SYTEM OF SYTEMS AND APPLICATION OF SCORE TO THE ADAPTIVE RESOURCE ALLOCATIONS

In the previous section, we demonstrated the engineering systems multiple-domain modeling approach to model and analyze the transforming health information SoS using the attributes of the process, architecture, and risk associated with the environment in the case of the Cambodian malaria surveillance system. The ES-MDM modeling framework [114] was applied and then the scoring approach was selected using the process, architecture, and environment to analyze the relative weights of the constituent systems. Similarly, the environmental condition can change over time, affecting the causations of several diseases such as malaria. Hence, in this chapter, we demonstrate the changes in all the aspects (process, architecture, and environment) employed in the scoring approach to capture or manage the transitional complexities in the SoS over time in the case of Cambodian malaria surveillance system. The transitions in the relative weights of the constituent systems are presented based on the results of transitional scoring of SoS and environmental risks calculated by the spatiotemporal model developed through the routinely aggregated data and remote sensing data captured by space satellites. We then demonstrate the adaptive resource allocation approach using the score as a case of application of the score.

## 5.1 Engineering systems multiple-domain modeling approach for the system

As described in the previous section, we employed the ES-MDM modeling frame for the system modeling and analyses. The ES-MDM was applied to examine the interactions of multiple domains, i.e., not just the process and architecture but also the environment of the system under investigation. We employed the case of the malaria surveillance system in Pailin province. We developed the DSM of the process and architecture domains before and after the reformation of the system. The interactions of multiple domains were then examined using DMM at the intersection of DSMs. We then calculated the relative weights of the constituent systems to investigate the transitional complexities caused by both the system reformation and continuous interactions of the environment by multiplying the three factors, i.e., the architectural, process, and environmental attributes of the SoS. The detail explanation of the score component is found in chapter 4. This time, we employed the spatiotemporal mathematical model of SMRs to estimate the environmental attributes of constituent systems at each time interval. We calculated them using the reported annual parasites incidence of malaria by the routine surveillance and environmental covariates captured from the space satellites as well as the non-environmental anthropogenic covariates, such as the status of the bed net distribution and reported artemisinin resistance in western Cambodia. The detailed description of model building and results can also be found in chapter 3. For interpretation and comparison at different time points, the calculated values were normalized in the range of 0–1 using the min–max normalization method at the same time point whereby the changes owing to the system transformation ($\Delta S_i$) can be measured. The values were then used to represent the size of a network node when the network models were plotted at each time point. The visualization of these network models was achieved by the R software (https://www.r-project.org). We calculated the score of the relative weights of the key constituent systems between 2010 and 2013. During this period, the national health information system was reformed. Although the reformation may have certain steps to be fully implemented rather than occurring at once before its completion, we calculated the score under the condition that the reformation was completed between 2010 and 2011.

## 5.2 Visualization of the scores and the application to the adaptive resource allocation of the scores for managing transitional complexities of health information system

### 5.2.1 Transitions in the relative weights of the constituent systems

Table 10 lists the scores of the relative weights of the key constituents in the malaria surveillance system at each interval of time. As shown in Figure 44, the relative weights of the constituent systems vary to a particular extent in accordance with the changes in the architecture and process of the system and environmental risks. After the reformation of the system between 2010 and 2011, there is a decentralization of the weights from the higher data-aggregating layer to the peripheral field practitioners, such as the HCs, and VMWs. In association with the changes in the malaria risk at the locations of the constituent systems, their relative weights are also changed accordingly. Hence, the changes in the relative weights are observed after the reformation of the system between 2010 and 2013. Notably, the magnitude of the changes is larger in the peripheral constituent systems, such as in the HCs and with VMWs, than in the central data-aggregating layers. These scores can be visualized in the complex network within the system, which may facilitate the common understanding and shared decision-making of the multiple stakeholders. Figure 45 shows the visualized network using the scores in each interval of the study period. This study is performed under the condition that the system is operating appropriately as documented. However, it is unlikely that the system maintains such a state continuously in most real-world conditions. The advantage of the ES-MDM modeling frame is that the model can flexibly consider these changes and deviations from the designated state of the system, which may be more beneficial for open systems with increased complexity because they are subject to the changes along with the manner in which they interact with the environment.

**Table 10 Scores of relative weights for key constituents**

| Constituents | Scores | | | |
|---|---|---|---|---|
| | 2010 | 2011 | 2012 | 2013 |
| CNM | 0.055 | 0.017 | 0.016 | 0.014 |
| DPHI/MoH (DPHI) | 0.025 | 0.056 | 0.053 | 0.044 |
| PHD Pailin | 0.088 | 0.093 | 0.088 | 0.073 |
| HOD Pailin | 0.655 | 0.816 | 0.775 | 0.643 |
| RH Pailin town | 0.618 | 0.850 | 0.808 | 0.670 |
| HC1 Suon Koma | 0.329 | 0.608 | 0.650 | 0.457 |
| HC1 VMWs (VMW1) | 0.070 | 0.318 | 0.339 | 0.239 |
| HC2 Ou Chra | 0.468 | 0.704 | 0.741 | 0.566 |
| HC2 VMWs (VMW2) | 0.099 | 0.368 | 0.387 | 0.296 |
| HC3 Phom Spong | 0.642 | 0.825 | 0.808 | 0.664 |
| HC3 VMWs (VMW3) | 0.136 | 0.431 | 0.422 | 0.347 |
| HC4 Psar Prum | 0.655 | 0.880 | 0.907 | 0.764 |
| HC4 VMWs (VMW4) | 0.214 | 0.447 | 0.460 | 0.388 |
| HC5 Phnom Preal | 0.790 | 1.000 | 1.000 | 0.923 |
| HC5 VMWs (VMW5) | 0.259 | 0.507 | 0.507 | 0.468 |
| HC6 Kracharb | 1.000 | 0.960 | 0.803 | 1.000 |
| HC6 VMWs (VMW6) | 0.327 | 0.487 | 0.407 | 0.507 |

MoH Ministry of Health, DPHI Department of Planning and Health Information at MoH, HC Health Center (HC1, ..., HC6), RH Referral Hospital, HOD Health Operational District, PHD Provincial Health Department, CNM National Center for Parasitology, Entomology and Malaria Control, VMW Village Malaria Worker.

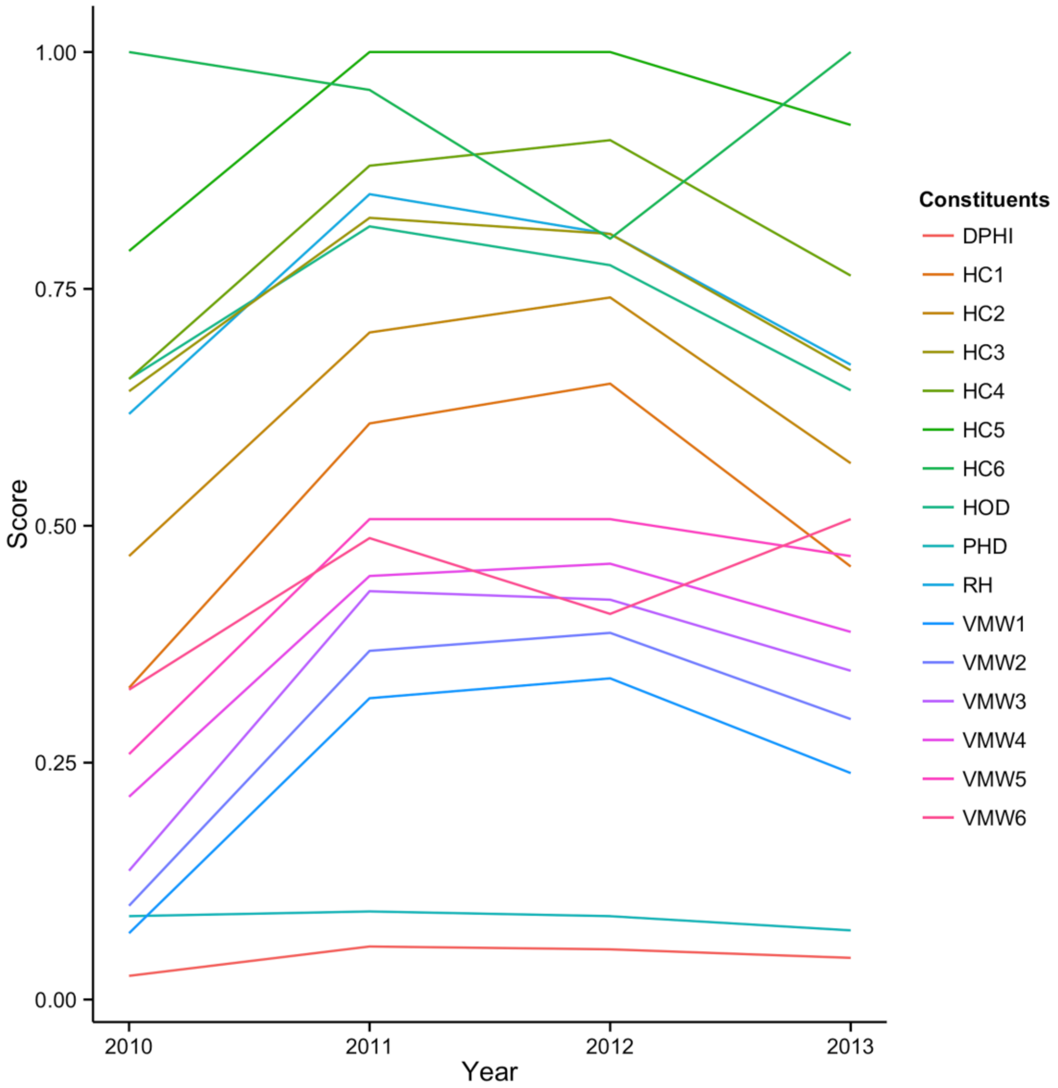



**Figure 44 Transition of the relative weights of the constituent systems**

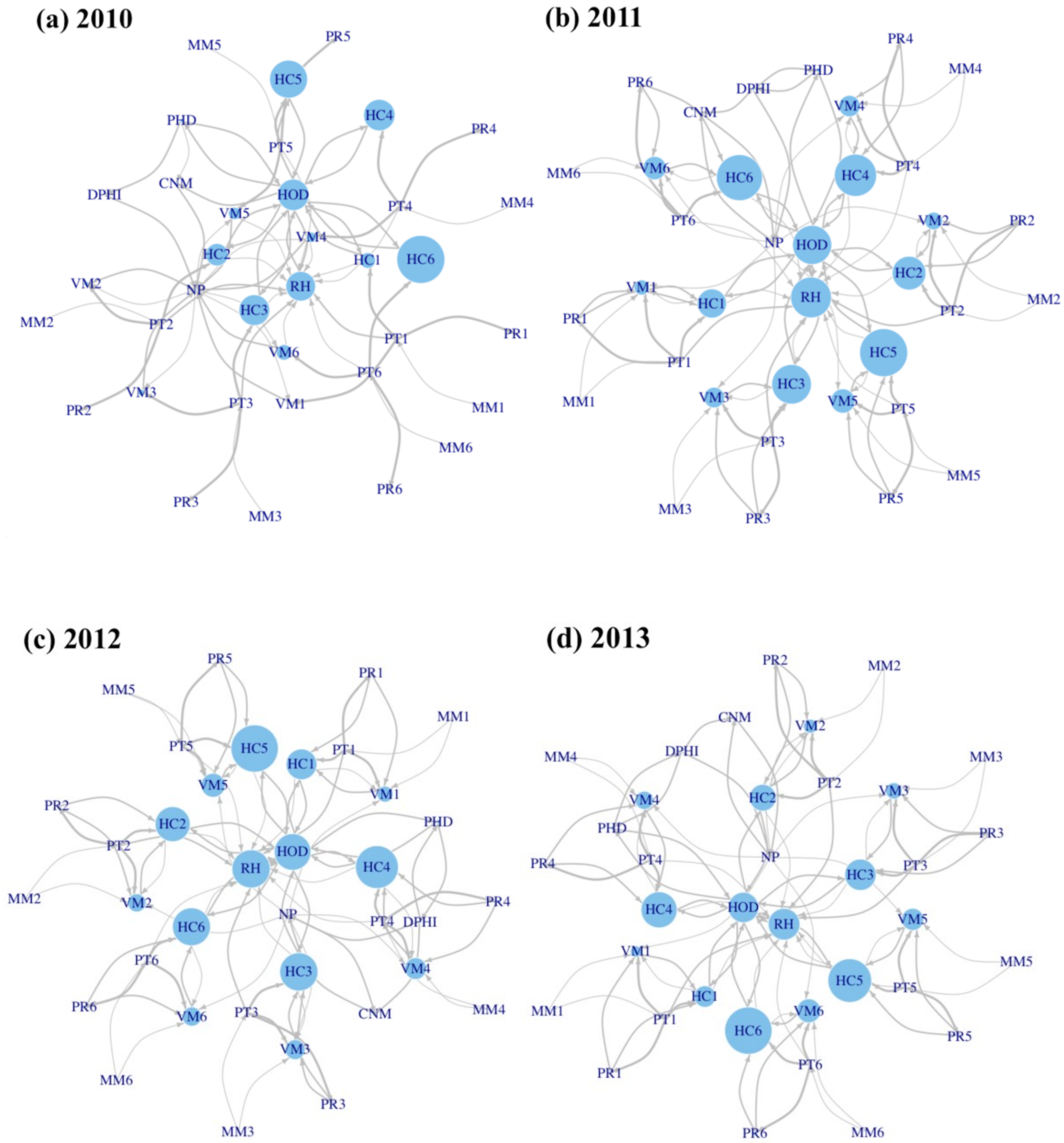



**Figure 45 Visualized network using the scores of the relative weights of the constituent systems in each interval of time**

### 5.2.2 Adaptive healthcare resource allocation using the score

In this section, we demonstrate the adaptive healthcare resource allocation using the scores of the relative weights of the constituent systems as an example of application. Because a score can capture the transition of the architectural, process, and environmental attributes over time, it can be used to solve the resource allocation problem for more achieving a tailored resource distribution over time by considering the changes occurring in each constituent system. Various interventions can be considered such as training for field practitioners, monitoring visits, and practical interventions. In general, the effect of intervention is correlated with the amount of resources allocated for the intervention. When there is reproducibility in the relationship between the effect of the intervention and amount of resources allocated, the effect of intervention can be mathematically modeled. Under the condition of limited resources, the resources need to be allocated to each constituent of the system to obtain the maximum effect of the intervention at each time interval, $t$, as expressed in (30), where $x_i$ is the amount of resource allocated to system constituent $i$ in $n$ constituent systems, $g_i(x_i)$ is the effect of the intervention, and $R_t$ is the available amount of resource at time interval $t$.

$$\begin{cases} max \sum_i g_i\,(x_i) \\ subject\ to \\ x_1 + x_2 + x_3\ +\ \ldots + x_n\ =\ R_t \end{cases} \qquad (30)$$

Assuming a nonlinear relationship between the effect of the intervention and amount of resource input, the former can be modeled as (31).

$$g(x) = a(1 - e^{bx}) \qquad (31)$$

We employed (31) for the reason that the model crosses 0 and assumes that the effect of the intervention reaches a plateau after adding a certain amount of resource. The effect at each constituent system can be modeled by multiplying the score of the constituent system and the modeled effect of intervention, as expressed in (32), where $S_i$ is the calculated score and $c_i$ is the specific factor that needs to be considered for constituent system $i$.

$$g_i(x_i) = S_i \cdot a(1 - e^{bx}) \cdot c_i \quad (32)$$

If the effect of the intervention can be transferred to the next year, the incremental gain of placing additional resources in year t can be calculated by (33).

$$g_{i_t}(x_{i_t}) = S_{i_t} \cdot a\{(1 - e^{bx_t}) - (1 - e^{bx_{t-1}})\} \cdot c_i \quad (33)$$

Assuming that approximately 50% of the maximum effect can be obtained by 2 units of resource input, 75% by 5 units, and 90% by 10 units, respectively, the parameters, a = 0.912 and b = -0.378, as shown in Figure 46. Under the condition that 50 units of resource are available from 2010 to 2013, the resource allocation problem for starting a new intervention is solved. We set specific factor $c_i = 0.5$ for remote places (HC 4–6 and VMW 4–6), considering that more resources are required to be delivered to these locations. In this study, we employed the dynamic programming approach to solve the resource allocation problem. Dynamic programming [135] is a mathematical method for effectively obtaining the appropriate optimal or equilibrium solution by treating the model as a sequence of decision processes. It is commonly used to solve the resource allocation problem. Table 11 lists the results of the resource allocation for constituents. The effect of the intervention was 3.582 in 2010, 1.880 in 2011, 0.597 in 2012, and 0.591 in 2013. The score guided the optimal resource allocation in accordance with the transitions in the architectural, process, and environmental attributes of the constituent systems, even under the condition that the reformation of the system occurred in the middle of the study period. One advantage of using ES-MDM to model the transitional condition of the health information SoS is that the model can be used to examine the expected effect caused by design phase, so that the stakeholders such as the model builder, planner, and decision-maker can visually track the examined changes in the SoS. The score can be calculated and used to solve the resource allocation problem, which enables more timely planning of the appropriate resources allocation immediately after the reformation of SoS.

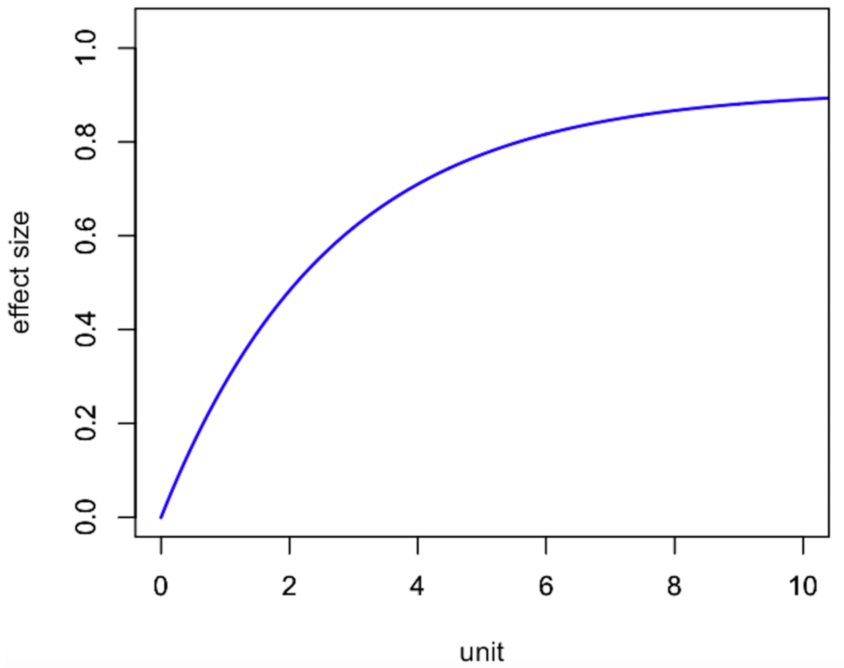



**Figure 46 Relationship between the healthcare resource input (unit) and the effect size of the intervention in the model used in the study**

**Table 11 Resource allocation at each constituent**

| Constituents | Resource allocation (units) | | | |
|---|---|---|---|---|
| | 2010 | 2011 | 2012 | 2013 |
| RH Pailin town | 5 | 2 | 3 | 2 |
| HC1 Suon Koma | 3 | 3 | 3 | 2 |
| HC1 VMWs (VMW1) | 0 | 5 | 2 | 5 |
| HC2 Ou Chra | 4 | 3 | 3 | 2 |
| HC2 VMWs (VMW2) | 0 | 5 | 3 | 4 |
| HC3 Phom Spong | 5 | 2 | 3 | 1 |
| HC3 VMWs (VMW3) | 1 | 4 | 3 | 4 |
| HC4 Psar Prum | 3 | 2 | 3 | 2 |
| HC4 VMWs (VMW4) | 1 | 3 | 2 | 2 |
| HC5 Phnom Preal | 4 | 2 | 3 | 2 |
| HC5 VMWs (VMW5) | 1 | 3 | 3 | 3 |
| HC6 Kracharb | 5 | 1 | 2 | 2 |
| HC6 VMWs (VMW6) | 2 | 2 | 2 | 2 |

HC Health Center (HC1, ..., HC6), RH Referral Hospital, HOD Health Operational District, VMW Village Malaria Worker.

# 6. DISCUSSION

## 6.1 Requirement verification traceability matrix

The Requirement Verification Traceability Matrix for the requirements need evaluations is shown in Table 12.

**Table 12 Requirement verification traceability matrix of the system**

| ID | Requirements | Verification method | Results |
|---|---|---|---|
| 1 | Estimate the risk of malaria under the low-to-moderate malaria transmission setting | Section 3.4: **Analysis**<br>Calibration plot and the model examination<br>Section 3.5: **Analysis**<br>Comparison of geocoded data and predicted data on the map<br>Section 3.7: **Analysis**<br>Observed and predicted uncertain range of SMR | Met |
| 1.1 | Visualize (Provide) information | Section 3.5/3.7: **Demonstration/Inspection**<br>Stakeholders interview | Met |
| 1.1.1 | Display estimated malaria risk | Section 3.5/3.7: **Inspection**<br>Inspection of visual representation of the map | Met |
| 1.1.2 | Display the simulation results | Section 3.5: **Demonstration**<br>Simulation and visual inspection of the results | Met |
| 1.2 | Support decision-making | Section 3.8: **Inspection**<br>Stakeholders interview | Met |
| 2 | Provide information without significant cost | Section 3.5/3.7: **Demonstration**<br>Providing results with data from publicly available sources | Met |
| 3 | Take dynamic situation and transformation of the SoS itself | Section 4.2/4.3: **Demonstration/Analysis**<br>Presenting multi-domain matrix of the routine malaria surveillance system<br>Validated by comparing the results from ABM simulation | Met |
| 3.1 | Reflect the changes in the architecture and the process before and after the reformation | Section 4.2: **Demonstration/Inspection**<br>Presenting multi-domain matrix of the system | Met |

| | | | |
|---|---|---|---|
| 3.2 | Reflect the spatiotemporal dynamics of environmental risk | Section 3.7/4.2/4.3: **Demonstration/Analysis** Estimate the spatiotemporal risk distribution Reflect environmental attributes of constituent systems for scoring relative weights, which was validated through the comparison with the results of ABM simulation | Met |
| 3.3 | Present the transitional changes of relative weights of constituent systems | Section 5.2: **Demonstration** Presenting transitional changes in relative weights of key constituent systems both numerically and visually | Met |
| 3.4 | Provide necessary information that aids in the continuous management of health information system | Section 5.2: **Demonstration** Demonstrated adaptive resource allocation as a case of application of the score | Met |

## 6.2 Validation of the proposed system

We aimed to develop the continuous management cycle of health information system through the iterative process of the mapping approach to estimate the spatial heterogeneity of disease burdens and its application for the modeling and analysis of the health information system for the purpose of better healthcare resource deployment while improving the quality of health information in the case of the malaria issue in western Cambodia. The specific objectives of study were: 1) To develop the spatial risk distribution model of malaria adjusted for the low-to-moderate transmission settings, considering environmental context disparities using routinely collected surveillance data in the health information system, remote sensing data, and publicly available data; 2) To develop an approach for modeling and analyzing health information SoS to understand the transitional complexities originated from the changes in the architecture, the process, and the surrounding environment for effective system management and; 3) To demonstrate the transitions of the relative weights of the constituent systems in the health information SoS and its application for the adaptive resource allocation for data quality interventions. Using the routinely collected surveillance data in conjunction with remote sensing data and publicly available containment status information, we developed the spatiotemporal model of disease risk of malaria and created the fine-scale risk maps. One strength of the proposed approach is that the source data were mostly from the publicly available source as such the fine-scale risk map could be created cost-effectively, which is in line with the requirement of cost-effective operation of the system under the low-to-moderate transmission settings. Furthermore, the results were validated through the comparison of geocoded data and estimated risk of each places and interviews with stakeholders. Hence, I believe that the proposed approach is an effective method to estimate the spatiotemporal risk distribution of malaria under the situation investigated in this study. While the supportive results were obtained in this study, it is important to understand that the API is reliable measure under the condition that the data collection is supported by the rigorous surveillance system. Besides, continuous effort of maintaining the surveillance system and provision of data to the public is critical to ensure that the proposed system will be sustainable.

The issue found in the stakeholders interview was reliability of reported data in Cambodia. Not only in Cambodia, the quality of health information continues to be of concern, particularly in situations with limited healthcare resources in some less industrialized countries. Health information system is a SoS that transforms over time through the interaction of changing environment and requirement of various stakeholders. Hence, it is important to understand the transitional conditions of this SoS in an ongoing manner for effective systems engineering. We applied the ES-MDM modeling framework to present the reformation of Cambodian malaria surveillance system. The model presented the changes in the architecture and the process. Not only presenting the structural changes, the model helped score the relative weights of constituents using the architectural, process, and the environmental attributes in the model to capture the transitional changes of key constituents as well as the structure of entire health information SoS. This approach was validated through the simulation test using ABM. Although the results were not conclusive, they provided further insights in how these scores can be used for SoS engineering. The model of the system presented in this study is merely the results of a single viewpoint of stakeholders. The level of detail and abstraction of the model may differ based on the viewpoint of the modelers. However, the model can readily accommodate changes to the system architecture and process, which is convenient for managing transitional complexities of an open system that transforms over time, such as ES. Despite the growing interest and needs for the governance of SoS engineering, there are few SoS standards in place today [136].

However, the effort toward standardization of SoS engineering may also provide the guidance for more standardized way.

Using the spatiotemporal risk model in combination with the modeled SoS by the ES-MDM framework, the transitional changes in the relative scores of key constituents were scored over the study period. The score could be applied to solve the adaptive resource allocation problem while considering the dynamics of environmental conditions and changes in the architecture and the process originated from the reformation. Not the single approach but the combination of risk mapping and modeling and analyses of SoS can provide the opportunity for more effective data utilization and feedback cycle. The proposed mapping approach can add more reciprocity among stakeholders than simply recording the aggregated case numbers. Through the continuous effort of more effective report-and-utilization cycles, data quality and reliability can be improved. Routine health information systems are in place in nearly every country and provide routinely collected full-coverage records on all level of health system service delivery [137]. However, these rich sources of data are regularly overlooked for evaluating various aspects of heath programs due to concerns of completeness, timeliness, representativeness, and accuracy. With rapidly growing technology, the capability to share the data – and harness its potential to generate knowledge rapidly and inform decision – can have transformative effects that improve the health by establishing the system called learning health care system [138]. The continuous improvement of health information in the health information SoS serves as the foundation in realizing the learning healthcare system.

## 6.3 Future work

While the proposed approach generate the several supportive results and implications continuous improvement cycle of health information by employing the mathematical risk mapping approach and modeling and analysis of transforming health information system, several limitations exist that need to be studied.

### 6.3.1 Limitation and future work for the risk mapping and engineering systems multiple-domain modeling approach

The limitations and future work for the risk mapping and engineering systems multiple-domain modeling approach were already discussed in the chapter 3 and 4, respectively.

### 6.3.2 Benefit of using the proposed approach

Although we demonstrated the transition of relative weights of key constituents in the health information system and the application of the scores for the adaptive resource allocation, it is required to examine the effectiveness of using the scores over the usual resource planning approach to understand more about the benefit of the proposed approach. One advantage of using ES-MDM to model the transitional condition of the health information is that the model can be applied to examine the expected effect caused by the changes in the architecture or process of SoS in the design phase, so that the stakeholders such as the model builder, planner, and decision-maker can visually track the examined changes in the SoS. Moreover, the model can flexibly reflect the real world situation, where deviations from designated state occur in most of the cases. However, over these qualitative advantages, the quantitative benefit such as the incremental gain of interventional effect and cost effectiveness to achieve the target effect need to be studied to facilitate the introduction of the new approach. Also, it might be important to consider numerous constraints and conditions when the approach will be applied in real cases. We demonstrated the resource allocation under the relatively simple conditions. As such, complementary method to cover this limitation such as sensitivity analyses or addition, other constraints or probabilistic approach to estimate the tailored resource allocation plan can be considered in future work.

### 6.3.3 Application to the other epidemiological issues or geographic areas

Strength of the proposed approach is utilization of existing dataset. Generally, most of such data sources are open to public. Hence, several more information can be added to the aggregated surveillance report in a real time manner. Since the proposed approach was demonstrated in the case of malaria surveillance in Cambodia, it is clear that the approach needs to be validated in other epidemiological issues or geographic areas. One area that we can consider this approach make contribution is the real time monitoring of emerging infectious disease such as Ebora [139] and recent severe acute respiratory syndrome (SARS) corona virus out break [140]. Also the application to non-communicable disease such as life-style disease needs to be elucidated considering future transition of disease structure globally. Recent advancement in the information technology enables the real-time collection of various kinds of data surrounding people. The recognition of Internet of Things (IoT) [141] is expanding and was actively introduced into the policy development in several countries [142]. Thus, we expect that the environmental context related to non-communicable disease could be understood using these advancing technologies such as atypical information surrounding people.

# 7. CONCLUSION

The health information system is a SoS that transforms over time in response to the changing environment and needs of the stakeholders. Though the reliability and timeliness of the data from the health information system play critical roles in ensuring sustainable access to healthcare services, the quality of the data continues to be of concern. In this thesis we developed the continuous management cycle of health information system through the iterative process of the mapping approach to estimate the spatial heterogeneity of disease burdens and its application for the modeling and analysis of the health information system for the purpose of better healthcare resource deployment while improving the quality of health information in the case of the malaria issue in western Cambodia. We developed the spatial risk distribution model of malaria adjusted for the low-to-moderate transmission settings, considering environmental context disparities using routinely collected surveillance data in the health information system, remote sensing data, and publicly available data. Next, we demonstrated the engineering systems multiple-domain modeling approach to capture the transitional complexities of the health information SoS for the effective system engineering. The engineering systems multiple domain modeling framework was applied to capture the reformation of the architecture and process of the Cambodian malaria surveillance system. The model could flexibly accommodate not only the changes resulting from the reformation of the system but also the deviations from the designated state, which could assist in reflecting practical real world situations and provide opportunities for more relevant analyses. Using the architectural, process, and environmental attributes of the constituent systems, the transitional changes in the relative weights were scored in each interval of time, which helped in understanding the transitional complexities of the health information SoS. These scores could be used to facilitate the healthcare resource allocation on a timely basis while causing the decision-maker to consider changes in the environment. This approach provided further steps for analyzing transitional conditions of the constituent systems as well as the entire health information SoS, whereby informed decision making for optimizing the continuous system management was facilitated. With this initial step of scoring the transitional changes in the constituent systems for continuous systems management, several remaining questions are identified that require further investigations. Further progress in the quality improvement of health information can be attained by the investigation of such work. The effort of continuous feedback cycle of the appropriate revisions of the modeling and provisions of the management actions in the healthcare system can contribute to providing sustained healthcare access to people.

# 8. PUBLICATON

This thesis has allowed the following scientific productions.

**Conference**

[1] **Okami S**, Kohtake N, (2015) Designing the GIS predicting regional malaria emdemicity in Cambodia. Esri Health and Human Services GIS Conference, Grand Hyatt Atlanta in Buckhead, Atlanta, Georgia.

[2] **Okami S**, Kohtake N (2017) Modeling and analysis of healthcare information system-of-systems for managing transitional complexities using engineering systems multiple domain matrix, Proc. IEEE International Systems Conference. P.688–695.

[3] **Okami S**, Kohtake N (2018) Managing health information system-of-systems by engineering systems multiple-domain modeling approach considering spatiotemporal dynamics. IEEE International Symposium on Systems Engineering (Accepted).

**Journal**

[1] **Okami S**, Kohtake N. (2016) Fine-scale mapping by spatial risk distribution modeling for regional malaria endemicity and its implications under the low-to-moderate transmission setting in western Cambodia. PLoS ONE 11 (7): e0158737.doi:10.1371/journal.pone.0158737.

[2] **Okami S**, Kohtake N. (2017) Spatiotemporal modeling for fine-scale maps of regional malaria endemicity and its implications for transitional complexities in a routine surveillance network in western Cambodia. Frontiers in Public Health 5: 262. doi:10.3389/fpubh.2017.00262

[3] **Okami S**, Kohtake N. (2017) Transitional complexity of health information system of systems: managing by the engineering systems multiple-domain modeling approach. IEEE Systems Journal. doi:10.1109/JSYST.2017.2778418

# 9. BIBLIOGRAPHY


[1]. World Health Organization (2004) World medicine situation. Geneva: World Health Organization, 2004, 61.

[2]. Frost L.J. and Reich M.R. (2008) Access: how do good health technologies get to poor people in poor countries?. Cambridge, Massachusetts: Harvard Center for Population and Development Studies, 2008, ISBN 978-0-674-03215-6.

[3]. UN Assembly (2015) Transforming our world: the 2030 agenda for sustainable development. 21 October 2015, A/RES/70/1, available at: http://www.refworld.org/docid/57b6e3e44.html.

[4]. Gottret P. and Schieber G. (2006) Health financing revisited: a practitioner's guide. Washington DC: The World Bank, 2006, 36.

[5]. Nakaya T. (2008) Spatial epidemiology in Geographic Information System environments. *J. Natl. Inst. Public Health* **57** (2): 99–116.

[6]. Wesolowski A., Eagle N., Tatem A.J., Smith D.L., Noor A.M., et al. (2012) Quantifying the impact of human mobility on malaria. *Science* **338**: 267–270.

[7]. Hay S.I., Snow R.I., Rogers D.J. (1998) From predicting mosquito habitat to malaria seasons using remotely sensed data: practice, problems and perspectives. *Parasitology Today* **14** (8): 306–313.

[8]. Shimodaira H., Nozaki M., Kwon Y., Kamimura N., Kaiho F. (2014) Analysis of adverse reaction in Kampo-medicines using JADER database of PMDA. *Jpn. J. Drug Inform.* **16** (1): 16–22.

[9]. World Health Organization (2000) Design and implementation of health information systems. Geneva, Switzerland: World Health Organization.

[10]. World Health Organization (2008) Framework and standards for country health information systems. Geneva, Switzerland: World Health Organization.

[11]. Hay S.I., Smith D.L., and Snow R.W. (2008) Measuring malaria endemicity from intense to interrupted transmission. *Lancet Infectious Diseases*, **8**, 369–378.

[12]. Mphatswe W., et al. (2012) Improving public health information: a data quality intervention in KwaZulu-Natal, South Africa. *Bull. World Health Org.*, **90**, 176–182.

[13]. Meier M.W. (1998) Architecting principles for system of systems. *Syst. Eng.*, 1, 267–284.

[14]. Bar-Yam Y. (1997) Dynamics of complex systems. Reading, MA, USA: Addison-Wesley.

[15]. Krishnan A., Nongkynrih B., Yadav K., Singh S., Gupta V. (2010) Evaluation of computerized health management information system for primary health care in rural India. *BMC Health Serv. Res.*, **10**. 310.

[16]. Maokola W., Willey B.A., Shirima K., Chemba M., Armstrong Shellenberg J.R., et al. (2011) Enhancing the routine health information system in rural southern Tanzania: successes, challenges, and lessons learned. *Trop. Med. Int. Health*, **16**(6), 721–730.

[17]. Mette S. and Morten O. (2012) Quality of cancer registry data: completeness of TNM staging and potential implications. *Clinical Epidemiology*, **4**, 1–3.

[18]. Hotchkiss D., Aqil A., Lippeveld T., and Mokooyo E. (2010) Evaluation of the performance of routine information system management (PRISM) framework: evidence from Uganda. *BMC Health Services Research*, **10**, 188.

[19]. Aqil A., Lippeveld T., and Hozumi D. (2009) PRISM framework: a paradigm shift for designing, strengthening and evaluating routine health information systems. *Health Policy and Planning*, **24**, 217–228.

[20]. Glèlè Ahanhanzo Y., Ouendo E.M., Kpozèhouen A., Levêque A., Makoutodé M., et al. (2015) Data quality assessment in the routine health information system: an application of the quality sampling in Berlin. *Health Policy and Planning*, **30**(7), 837–843.

[21]. World Health Organization (2015) Fact sheet N° 94. Geneva: World Health Organization.

[22]. WHO-UNICEF Child Health Epidemiology Reference Group (2013) Causes of deaths among children under 5 years, 2011. Geneva: World Health Organization.

[23]. World Health Organization (2007) Malaria elimination. A field manual for low and endemic countries. Geneva: World Health Organization.

[24]. Tatem A.J., Smith D.L., Gething P.W., Kabaria C.W., Snow R.W., Hay S.I. (2010) Ranking of elimination feasibility between malaria-endemic countries. *Lancet* **376**: 1579–1591.

[25]. National Center for Parasitology, Entomology and Malaria Control (2011) Strategic plan for elimination of malaria in Cambodia 2011-2025. Phnom Penh, Cambodia: Ministry of Health.

[26]. Maude R.J., Nguon C., Ly P., Bunkea T., Ngor P., et al. (2014) Spatial and temporal epidemiology of clinical malaria in Cambodia 2004-2013 *Malaria Journal* **13**: 385.

[27]. World Health Organization (2014) World malaria report 2014 Geneva: World Health Organization.

[28]. World Health Organization (2013) Emergency response to artemisinin resistance in the Greater Mekong subregion. Regional framework for action 2013-2015. Geneva: World Health Organization.

[29]. World Health Organization (2015) Guideline for the treatment of malaria. Third edition. Geneva: World Health Organization.

[30]. World Health Organization (2015) Q&A on artemisinin resistance. Geneva: World Health Organization.

[31]. World Health Organization (2015) Update on artemisinin and ACT resisitance – September 2015. Geneva: World Health Organization.

[32]. Ashley E.A., Dhorda M., Fairhurst R.M., Amaratunga C., Lim P., et al. (2014) Spread of artemisinin resistance in *Plasmodium falciparum* malaria. *New Engl J Med* **371**: 411–423.

[33]. Bosman P., Stassijns J., Nackers F., Canier L., Kim N., et al. (2014) *Plasmodium* prevalence and artemisinin-resistant falciparum malaria in Preah Vihear province, Cambodia: a cross-sectional population-based study. *Malaria Journal* **13**: 394.

[34]. Amaratunga C., Sreng S., Suon, S., Phelps E.S., Stepniewska K., et al. (2012) Artemisinin-resistant *Plasmodium falciparum* in Pursat province, western Cambodia: a parasite clearance rate study. *Lancet Infect Dis.* **12**(11): 851–858.

[35]. Laurent B.S., Miller B., Burton T.A., Amaratunga C., Men S., et al. (2015) Artemisinin-resistant *Plasmodium falciparum* clinical isolates can infect diverse mosquito vectors of Southeast Asia and Africa. *Nature Communications* **6**: 8614 doi: 10.1038/ncomms9614.

[36]. Leang R., Barrette A., Bouth D.M., Menard D., Abdur R., et al. (2013) Efficacy of Dihydroartemisinin-piperaquine for treatment of uncomplicated *Plasmodium*

*falciparum* and *Plasmodium vivax* in Cambodia 2008 to 2010. *Antimicrobial Agents and Chemotherapy* **57**: 818-826.

[37]. Leang R., Taylor W.R.J., Bouth D.M., Song L., Tarning, J., et al. (2015) Evidence of Plasmodium falciparum malaria multidrug resistance to artemisinin and piperaquine in western Cambodia: Dihydroartemisinin-piperaquine open-label multicenter clinical assessment. *Antimicrobial Agents and Chemotherapy* **59**(8): 4719–4726.

[38]. National Center for Parasitology, Entomology and Malaria Control (2014) Benefits to national malaria programs from regional support: the Cambodia case Phnom Penh, Cambodia: Ministry of Health.

[39]. Tabernero P., Fernández F.M., Green M., Guerin P.J., Newton P.N. (2014) Mind the gaps – the epidemiology of poor-quality antimalarials in the malarious world – analysis of the WorldWide Antimalarial Resistance Network database. *Malaria Journal* **13**: 139.

[40]. Guyant P., Canavati S.E., Chea N., Ly P., Whittaker M.A., Roca-Felter A. et al. (2015) Malaria and the mobile and migrant population in Cambodia: a population movement framework to inform strategies for malaria control and elimination. *Malaria Journal* **14**: 252.

[41]. Hoyer S., Nguon S., Kim S., Habib N., Khim N., Sum N., et al. (2012) Focused screening and treatment (FSAT): A PCR-based strategy to detect malaria parasite carriers and contain drug resistant *P. falciparum*, Pailin, Cambodia. *PLOS ONE* **7**(10): e45797.

[42]. Cox J., Soley L.D., Bunkea T., Sovannaroth S., Ty K.S., et al. (2014) Evaluation of community-based systems for the surveillance of day three-positive *Plasmodium falciparum* cases in western Cambodia. *Malaria Journal* **13**: 282.

[43]. Lwin K.M., Imwong M., Suangkanarat P., Jeeyapant A., Vihokhern B., et al. (2015) Elimination of *Plasmodium falciparum* in an area of multi-drug resistance. *Malaria Journal* **14**: 319.

[44]. Ministry of Health, Cambodia (2008) Heath information strategic plan 2008 – 2015. Phnom Penh: Ministry of Health, Cambodia.

[45]. Ministry of Health, Cambodia (2005) Health information system guideline first edition. Phnom Penh: Ministry of Health, Cambodia.

[46]. Ministry of Health, Cambodia (2011) Bulletin 2011: Cambodia web-based health information system. Phnom Penh: Ministry of Health, Cambodia.

[47]. Okell LC, Bousema T, Griffin JT, Ouédraogo AL, Ghani AC. (2012) Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nature Communications* **3**: 1237.

[48]. Sabot O., Tulloch J., Basu S., Dyckman W., Moonasar D., Moonen B. (2009) Getting to zero. *A Plospectus on Malaria Elimination* 19–39.

[49]. Aday L.A. and Andersen R. (1974) A framework for the study of access to medical care. *Health Services Research*, **9**: 208–220.

[50]. Hanson K., Ranson M.K., Oliveira-Cruz V., and Mills A. (2003) Expanding access to priority health interventions: a framework for understanding the constraints to scaling-up. *Journal of International Development*, **15**: 1–14.

[51]. Martin C., Curtis B., Fraser C., Sharp B.. (2002) The use of GIS-based malaria information system for malaria research and control in South Africa. *Health & Place* **8**: 227–236.

[52]. Clements A.C., Reid H.L., Kelly G.C., Hay S.I. (2013) Further shrinking the malaria map: how can geospatial science help to achieve malaria elimination? *Lancet Infect Dis.* **13**(8): 709–718.

[53]. Hay S.I., Guerra C.A., Gething P.W., Patil A.P., Tatem A.J., et al. (2009) A world malaria map: Plasmodium falciparum endemicity in 2007. *PLOS Medicine* **6**(3): e1000048.

[54]. Gething P.W., Patil A.P., Smith D.L., Guerra C.A., Elyazar I.R., et al. (2011) A new world malaria map: Plasmodium falciparum endemicity in 2010. *Malaria Journal* **10**: 378.

[55]. Patil A.P., Gething P.W., Piel F.B., Hay S.I. (2011) Bayesian geostaticstics in health cartography: the perspective of malaria. *Trends in Parasitology* **27**(6): 245–252.

[56]. Nihei N., Hashida Y., Kobayashi M., Ishii A. (2002) Analysis of malaria endemic areas on the Indochina Peninsula using remote sensing. *Jpn J Infect Dis.* **55**(5): 160–166.

[57]. Cohen J.M., Ernst K.C., Lindblade K.A., Vulule J.M., John C.C., et al. (2010) Local topographic wetness indices predict household malaria risk better than land-use and land-cover in the western Kenya highlands. *Malaria Journal* **9**: 328.

[58]. MacCann R.S., Messina J.P., MacFarlane D.W., Bayoh M.N., Vulule J.M., et al. (2014) Modeling larval malaria vector habitat locations using landscape features and cumulative precipitation measures. *International Journal of Health Geographics* **13**: 17.

[59]. Cianci D., Hartemink N., Justicia I. (2015) Modeling the potential spatial distribution of mosquito species using three different techniques. *International Journal of Health Geographics* **14**: 10.

[60]. Craig M.H., Snow R.W., Sueur D.I. (1999) A climate-based distribution model of malaria transmission in Sub-Saharan Africa. *Parasitology Today* **15**(3): 105–111.

[61]. Oloukoi G., Bob U., Jaggernath J. Perception and trends of associate health risks with seasonal climate variation in Oke-Ogun region, Nigeria. *Health & Place* **25**: 47–55.

[62]. Cohen J.M., Dlamini S., Novotny J.M., Kandula D., Kunene S., et al. (2013) Rapid case-based mapping of seasonal malaria transmission risk for strategic elimination planning in Swaziland. *Malaria Journal* **12**: 61.

[63]. Sturrock HJW, Cohen JM, Keil P, Tatem AJ, Menach AL, et al. (2014) Fine-scale risk mapping from routine aggregated case data. *Malaria Journal* **13**: 421.

[64]. Macdonald, G., (1957) The epidemiology and control of malaria. London: Oxford University press.

[65]. Dietz K., Molineaux L., Thomas A. (1974) A malaria model tested in the African savannah. *Bulletin of the World Health Organization* **50**: 347–357.

[66]. Chikodzi D. (2013) Spatial modeling of malaria risk zones using environmental, anthropogenic variables and geographical information systems techniques. *Journal of Geoscience and Geomatics* **1**(1): 8–14.

[67]. Ellis E.C., Ramankutty N. (2008) Putting people in the map: anthropogenic biomes of the world. *Front Ecol. Environ.* **6**(8): 439–447.

[68]. Smith D.L., Smith T.A., Hay S.I. (2009) Measuring malaria for elimination. *A Plospectus on Malaria Elimination* 108–126.

[69]. Hermens L.A. and Schlimmer J.C. (1994) A machine-learning apprentice for the completion of repetitive forms. *IEEE Expert*. **9**(1): 28–33.

[70]. Chen H., Chen K., Conway N., Hellerstein J.M., and Parikh T.S. (2011) Usher: improving data quality with dynamic forms. *IEEE Trans. Knowl. Data Eng.*, **23**(8), 1138–1153.

[71]. Dasu T. and Johnson T. (2003) Exploratory data mining and data cleaning (Wiley series in probability and statistics). Hoboken, NJ, USA: Wiley.

[72]. Graves R.M., Fowler F.J., Couper M.P., Lepkowski J.M., Singer E., and Tourangeau R. (2004) Survey Methodology. Hoboken, NJ, USA: Wiley.

[73]. Day S., Fayers P., and Harvey D. (1998) Double data entry: what value, what price? *Controlled Clin. Trials*, **19**: 15–24.

[74]. den Broeck J.V., Mackay M., Mpontshane N., Lubeya A.K.K., Chhagan M., and Bennish M.L. (2007) Maintaining data integrity in a rural clinical trial. *Clin Trials*, **4**: 572–582.

[75]. Grigoroudis E. and Phillis Y.A. (2013) Modeling healthcare system-of-systems: a mathematical programming approach. *IEEE Syst. J.*, **7**(4): 571–580.

[76]. Harrison W.K. (2016) The role of graph theory in system of systems engineering. *IEEE Access*, **4**: 1716–1742.

[77]. Center for Disease Control and Prevention. (2006) Principles of epidemiology in public health practice, third edition. An introduction to applied epidemiology and biostatistics. Atlanta, Georgia: CDC.

[78]. National Center for Parasitology, Entomology and Malaria Control. (2011) Cambodia malaria bulletin, December 2011. Phnom Penh, Cambodia: Ministry of Health.

[79]. National Center for Parasitology, Entomology and Malaria Control. (2013) Cambodia malaria bulletin, December 2013. Phnom Penh, Cambodia: Ministry of Health.

[80]. Mellor, S. (2013) Moving towards malaria elimination: developing innovative tools for malaria surveillance in Cambodia. available at: *www.malariaconsortium.org/pages/learning-paper.htm*

[81]. Tango T., Yokoyama T., Takahashi K. (2007) Introduction to spatial epidemiology. Tokyo, Japan: Asakura publishing.

[82]. National Center for Parasitology, Entomology and Malaria Control. (2010) Cambodia malaria survey 2010. Phnom Penh, Cambodia: Ministry of Health.

[83]. Lowe R., Chirombo J., Tompkins A.M. (2013) Relative importance of climatic, geographic and socio-economic determinants of malaria in Malawi. *Malaria Journal* **12**: 416.

[84]. Jenski D.E. and Wu J. (1996) The modifiable areal unit problem and implication for landscape ecology. *Landscape Ecology*, **11**: 129–140.

[85]. Cheng T. and Adepeju M. (2014) Modifiable temporal unit problem (MTUP) and its effect on space-time cluster detection. *PLOS ONE*, **9**(6): e100465.

[86]. Martuzzi M., Elliott P. (1996) Empirical Bayese estimation of small area prevalence of non-rare conditions. *Statistics in Medicine.* **15**: 1867–1873.

[87]. Mayer D.J., Tachikawa T., Abrams M., Crippen R., Krieger T., et al. (2012) Summary of the validation of the second version of the ASTER GDEM. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.* XXXIX-B4: 291–293.

[88]. Nmor J.C., Sunahara T., Goto K., Futami K., Sonye G., et al. (2013) Topographic models for predicting malaria vector breeding habitats: potential tools for vector control managers. *Parasites & Vectors* **6**: 14.

[89]. Gething P.W., Boeckel T.P.V., Smith D.L., Guerra C.A., Patil A.P., et al. (2011) Modeling the global constraints on transmission of *Plasmodium falciparum* and *P. vivax*. *Parasites & Vectors* **4**: 92.

[90]. Moyes C.L., Temperley W.H., Henry A.J., Burgert C.R., Hay S.I. (2013) Providing open acess data online to advance malaria research and control. *Malaria Journal* **12**: 161.

[91]. Tatem A.J., Guerra C.A., Kabaria C.W., Noor A.M., Hay S.I. (2008) Human population, urban settlement patterns and their impact on *Plasmodium falciparum* malaria endemicity. *Malaria Journal* **7**: 218.

[92]. Tatem A.J., Gething P.W., Smith D.L., Hay S.I. (2013) Urbanization and the global malaria recession. *Malaria Journal* **12**: 133.

[93]. Kurshcke J.K. (2014) *Doing Bayesian data analysis, 2nd edition.* Academic Press – Elsevier.

[94]. Brooks S.P., Gelman A. (1998) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**(4): 434-455.

[95]. Médecins Sans Frontières (2005) Unpublished dataset.

[96]. Incadona S., Vong S., Chiv L., Lim P., Nhem S., et al. (2007) Large-scale malaria survey in Cambodia: Novel insights on species distribution and risk factors. *Malaria Journal* **6**: 37.

[97]. Altman D.G. (1991) *Practical statistics for medical research.* London: Capman & Hall.

[98]. Welch B.L. (1938) The significance of the difference between two means when the population variances are unequal. *Biometrica* **29**: 350-362.

[99]. Carpenter J., Bithell J., et al. (2000) Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine* 19: 1141-1164.

[100]. World Health Organization. (2014) From malaria control to malaria elimination: a manual for elimination scenario planning. Geneva: World Health Organization.

[101]. Canavati S.E., Lawpoolsri S., Quintero C.E., Nguon C., Ly P., Pukrittayakamee S., et al. (2016) Village malaria worker performance key to the elimination of artemisinin-resistant malaria: a western Cambodia health system assessment. *Malaria Journal*, **15**: 282.

[102]. Bonacich P. (1972) Factoring and weighting approaches to status scores and clique identification. *J. Math. Sociol.*, **2**: 113–120.

[103]. Okami S., Kohtake N. (2016) Fine-scale mapping by spatial risk distribution modeling for regional malaria endemicity and its implications under the low-to-moderate transmission setting in western Cambodia. *PLOS ONE*, **11**: e0158737.

[104]. Trape J-F, Lefebvre-Zante E., Legros F., Ndiaye G., Bouganali H., et al. (1992) Vector density gradients and the epidemiology of urban malaria in Dakar. *American Journal of Tropical Medicine and Hygiene* **47**: 181–189.

[105]. Ghebreyesus T.A., Haile M., Witten K.H., Getachew A., Yohannes A.M., et al. (1999) Incidence of malaria among children living near dams in northern Ethiopia: community based incidence survey. *British Medical Journal* **319**: 663–666.

[106]. Carter R., Mendis K.N., Roberts D., et al. (2000) Spatial targeting of interventions against malaria. *Bulletin of the World Health Organization* **78**(12): 1401–1411.

[107]. National Center for Parasitology, Entomology and Malaria Control. (2005) Report of the Cambodia national malaria baseline survey 2004. Phnom Penh, Cambodia: Ministry of Health.

[108]. National Center for Parasitology, Entomology and Malaria Control. (2007) Cambodia malaria survey 2007 report. Phnom Penh, Cambodia: Ministry of Health.

[109]. Cohen J.M., Ernst K.C., Lindblade K.A., Vulule J.M., John C.C., Wilson M.L. (2008) Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western Kenyan highlands. *Malaria Journal* **7**: 40.

[110]. Targett GA, Yeung S, Tanner M. (2009) Identifying the gaps – what we need to know. *A Prospectus on Malaria Elimination.*

[111]. Bhatt S., Weiss D.J., Mappin B., Dalrymple U., Cameron E., Bisanzio D., et al. (2017) Coverage and system efficiencies of insecticide-treated nets in Africa from 2000 to 2017. *eLife*, **4**: e09672.

[112]. Baird J.K., Velcha N., Duparc S., White N.J., Price R.N. (2016) Diagnosis and treatment of *Plasmodium vivax* malaria. *Am. J. Trop. Med. Hyg.*, **95**(Suppl 6): 35–51.

[113]. Reid H.L., Haque U., Roy S., Islam N., Clements A.C.A. (2012) Characterizing the spatial and temporal variation of malaria incidence in Bangladesh. *Malaria Journal*, **11**: 170.

[114]. Bartolomei J.E., Hastings D.E., de Neufville R., and Rhodes D.H. (2012) Engineering systems multiple-domain matrix: an organizing framework for modeling large-scale complex systems. *Syst. Eng.*, **15**: 41–61.

[115]. Annear P.L., et al. (2015) The kingdom of Cambodia health system review. *Health Syst. Transition*, **5**(2).

[116]. Yasuoka J., et al. (2010) Assessing the quality of service of village malaria worker to strengthen community-based malaria control in Cambodia. *Malaria Journal*, **9**: 109.

[117]. Yasuoka J., et al. (2012) Scale-up of community-based malaria control can be achieved without degrading community health workers' service quality: the village malaria worker project in Cambodia. *Malaria Journal*, **11**: 4.

[118]. Boardman J. and Sauser B. (2006) System of systems: The meaning of of. *Proc. IEEE/SMC Int. Conf. Syst. Syst. Eng.*, Apr. p. 6.

[119]. Baldwin W.C. and Sauser B. (2009) Modeling the characteristics of system of systems. Proc. *IEEE Int. Conf. SoSE*, May–Jun. pp. 1–6.

[120]. ESD. (2008) Engineering systems strategic plan. Cambridge, MA, USA: Eng. Syst. Division, Massachusetts Inst. Technol.

[121]. de Weck O.L., Roos D., and Magee C.L. (2011) Engineering systems: meeting human needs in a complex technological world. Cambridge, MA, USA: Massachusetts Inst. Technol.

[122]. Bartolomei J.E., Cokus M., Dahlgren J., de Neufville R., Maldonaldo D., and Wilds J. (2007) Analysis and applications of design structure matrix, domain mapping matrix and engineering system matrix frameworks. Cambridge, MA, USA: Massachusetts Inst. Technol.

[123]. Browning T.R. (2001) Applying the design structure matrix to system decomposition and integration problems: a review and new directions. *IEEE Trans. Eng. Manage.*, **48**(3): 292–306.

[124]. Ulrich K.T. and Eppinger S.D. (2003) Product design and development, 3rd ed. Boston, MA, USA: MacGraw-Hill/Irwin.

[125]. Danilovic M. and Browning T.R. (2007) Managing complex product development projects with design structure matrices. *Int. J. Manage.*, **205**: 300–314.

[126]. Maurer M. and Linderman U. (2008) The application of the multiple-domain matrix: considering multiple domains and dependency types in complex product design. *Proc. IEEE Int. Conf. Syst. Man. Cybern.*, 2487–2493.

[127]. Bavelas A. (1950) Communication patterns in task-oriented groups. *J. Accounts. Soc. Amer.*, **22**: 725–730.

[128]. Bonabeau E. (2002) Agent-based modeling: methods and techniques for simulating human systems. *Proc. Nat. Acad. Sci.*, **99**: 7280–7287.

[129]. Epstein J.M. and Axtell R.L. (1996) Growing artificial societies: social science from the bottom up. Cambridge, MA, USA: MIT Press.

[130]. Wilsensky U. (1999) NetLogo, center connected learn. *Comput.-based model.*, Northwestern Univ. Evanston, IL, USA. [Online]. Available at: http://ccl.northwestern.edu/netlogo/

[131]. Bernardo J.M. and Smith A.F. (2000) Bayesian Theory (Wiley series in probability and statistics). Hoboken, NJ, USA; Wiley.

[132]. Bernd H., Marcus K., and Mathias K. (2007) How to measure data quality? A metric based approach. *Proc. 28th Int. Conf. Inf. Syst.*, Montreal, QC, Canada.

[133]. Xia C., Ding S., Wang C., Wang J., and Chen Z. (2017) Risk analysis and enhancement of cooperation yielded by the individual reputation in the spatial public goods game. *IEEE Syst. J.*, **11**(3), 1516–1525.

[134]. Wang C., Wang L., Wang J., S. Sun, and C. Xia. (2017) Inferring the reputation enhances the public goods game on interdependent lattices. *Appl. Math. Comput.*, **15**: 18–29.

[135]. Bellman R. (1954) The theory of dynamic programming. Bull. Amer. Math. Soc., **60**(66): 503–515.

[136]. Dahmann J., Roedler G. (2016) Moving towards standardization for system of systems engineering. *Proc. 11th IEEE System of Systems Engineering Conference 2016*, 1–6.

[137]. Wagenaar B.H., Sherr K., Fernandes Q., and Wagenaar A.C. (2015) Using routine health information systems for well-designed health evaluations in low- and middle-income countries. *Health Policy and Planning*, 1–7.

[138]. Friedman C., Rubin J., Brown J., Buntin M., Corn M., Etheredge L., et al. (2015) Toward a science of learning systems: a research agenda for the high-functioning learning health system. *J. Am. Med. Med. Inform. Assoc.*, **22**: 43–50.

[139]. Center for Diease Control and Prevention. (2016) 2014 Ebora outbreak in west Africa [Online]. Available: http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/.

[140]. Center for Disease Congtrol and Prevention. (2004) Fact sheet for SARS patients and their close contacts. Atlanta, Georgea: CDC.

[141]. Ashton K. (2009) That ‘Internet of Things’ Thing. *RFID Journal* 22 June. New York: RFID Journal LLC.

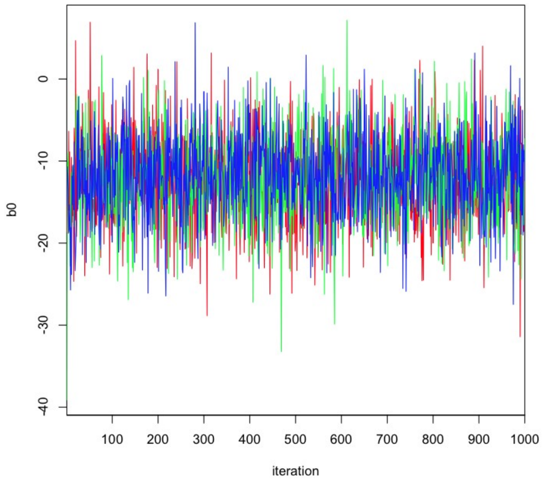
[142]. Kagermann H., Wahlster W., Helbig J., et al. (2013) Recommendations for implementing the strategic initiative Industrie 4.0: Final report of the Industrie 4.0 Working Group. Germany: The industry-science research alliance.

*Referenced website URLs are available as of May 2018.

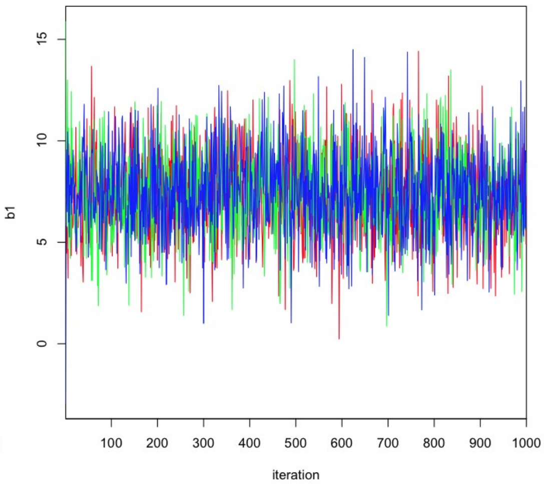
# 10. APPENDIX

Following figures are the trace plots of the Bayesian modeling frame to examine the convergence of each parameter.

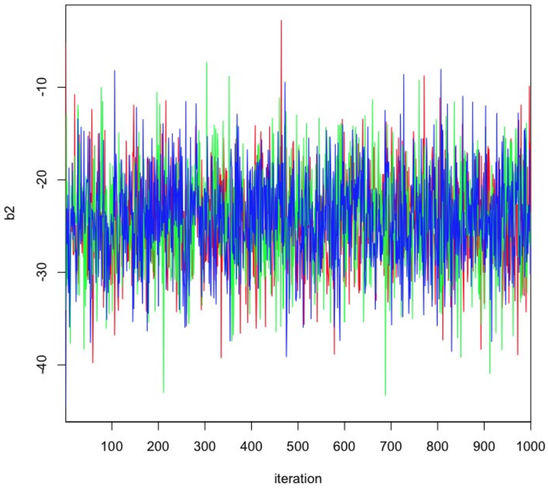
**A**



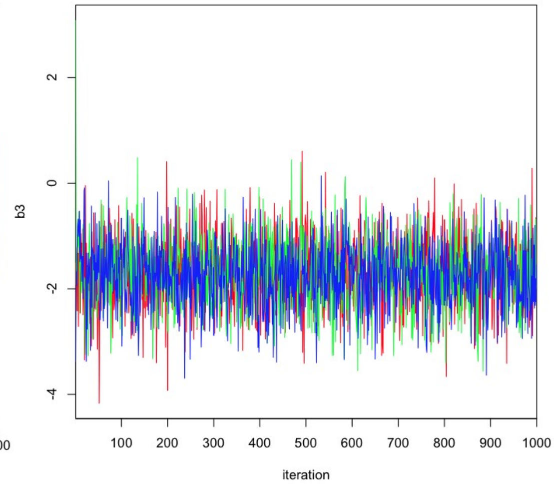

**B**



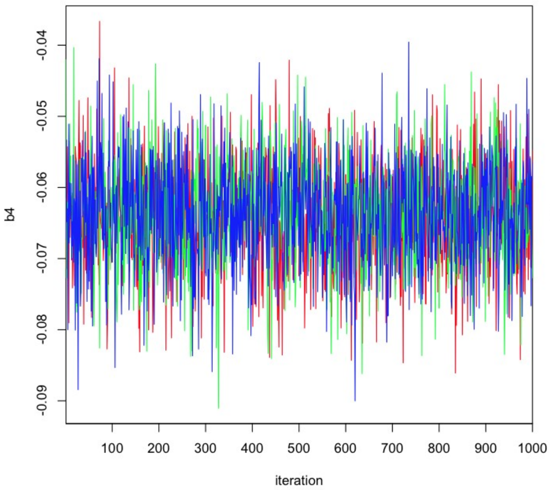

**C**



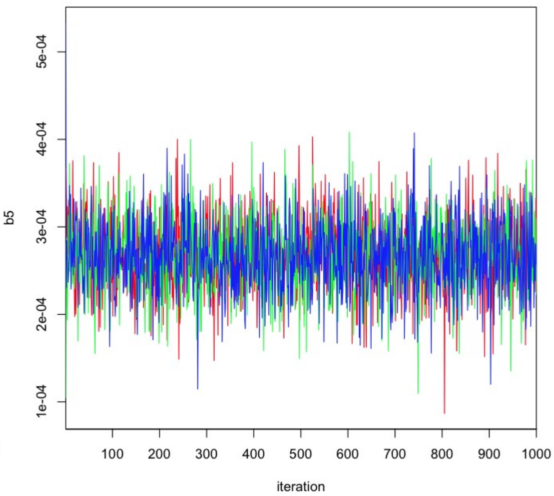

**D**




**E**




**F**



**G**

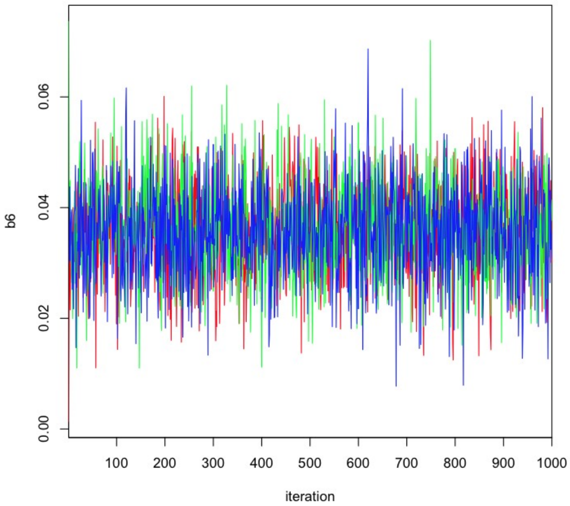
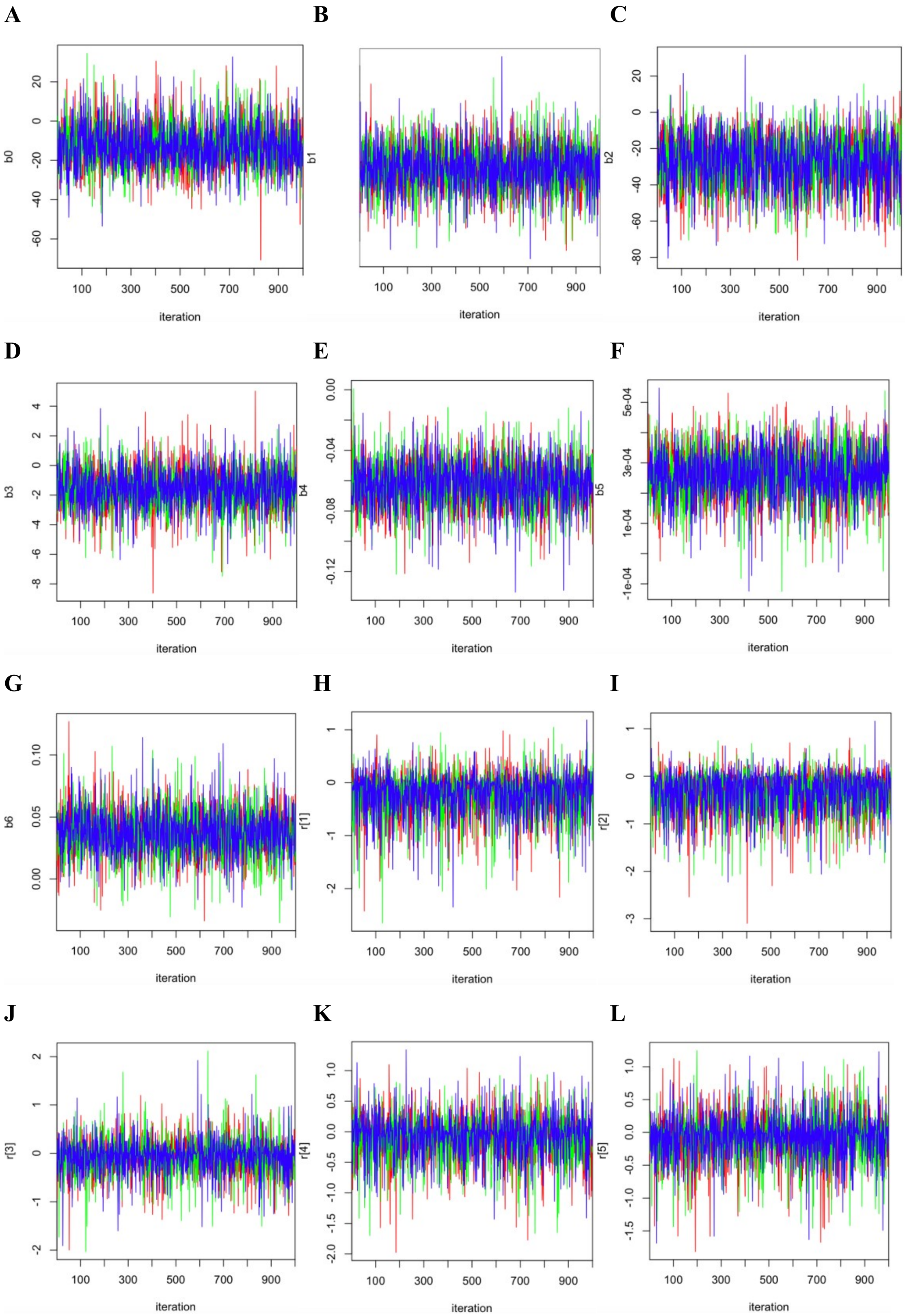
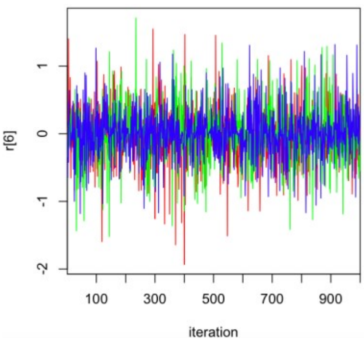



**Figure 47 Trace plot of the parameter in the Bayesian modeling for cross prediction for fine-scale malaria risk**

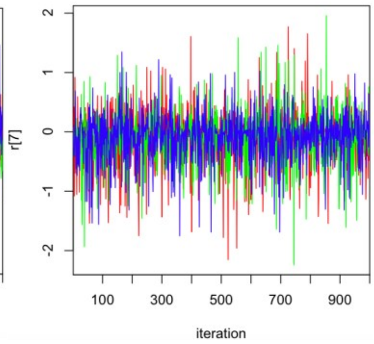
Parameter distributions for (A): Intercept, (B): NDVI, (C): NDWI, (D): TWI, (E): LLIN, (F): Temperature and (G): TF
NDWI, Normalized difference water index; NDVI, Normalized difference vegetation index; TWI, Topographical wetness index; LLIN, Long lasting insecticide-treated net


A
B
C
D
E
F
G
H
I
J
K
L
b0
b1
b2
b3
b4
b5
b6
r[1]
r[2]
r[3]
r[4]
r[5]
iteration

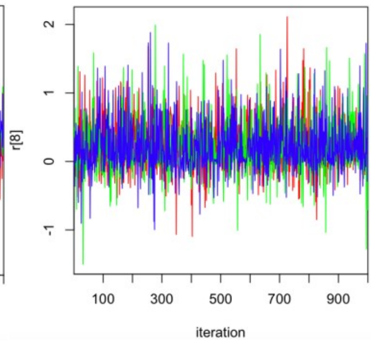
**M**



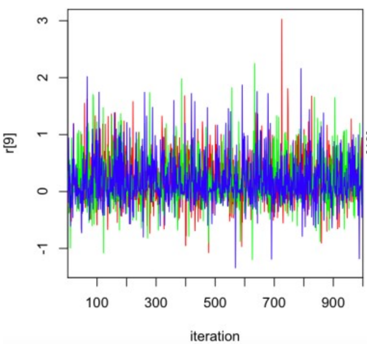

**N**



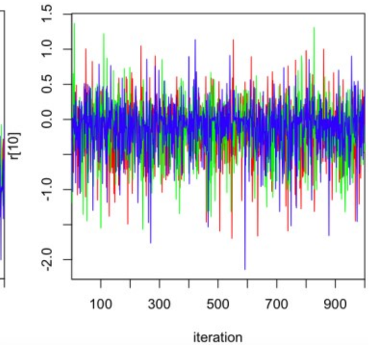

**O**



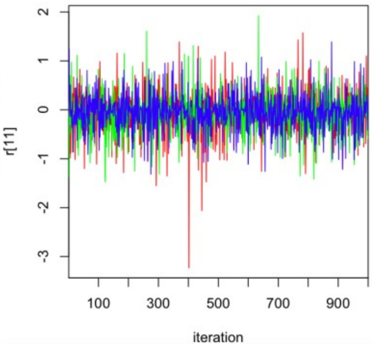

**P**



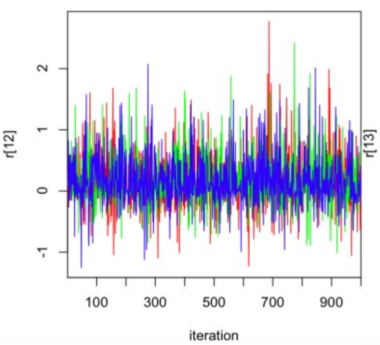

**Q**



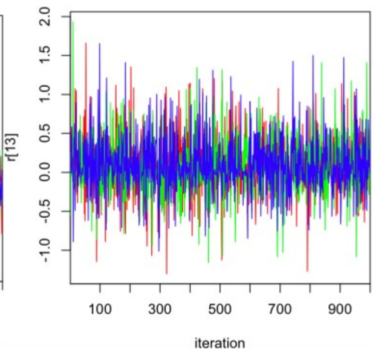

**R**



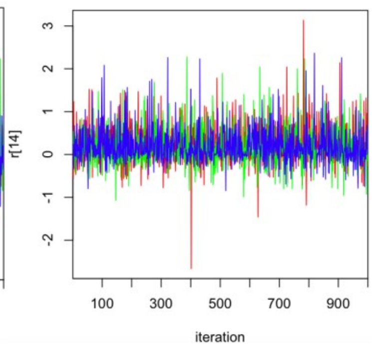

**S**



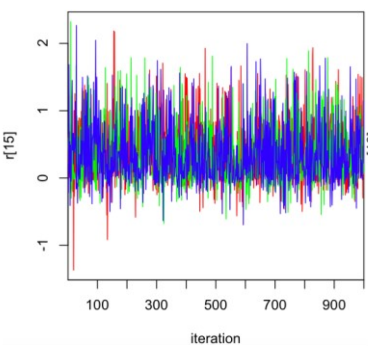

**T**



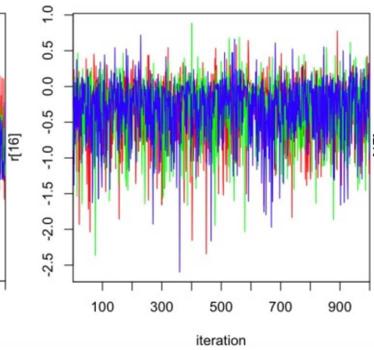

**U**



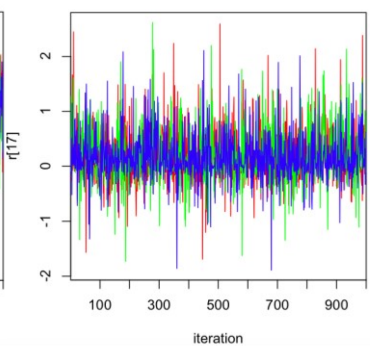

**V**




**W**




**X**



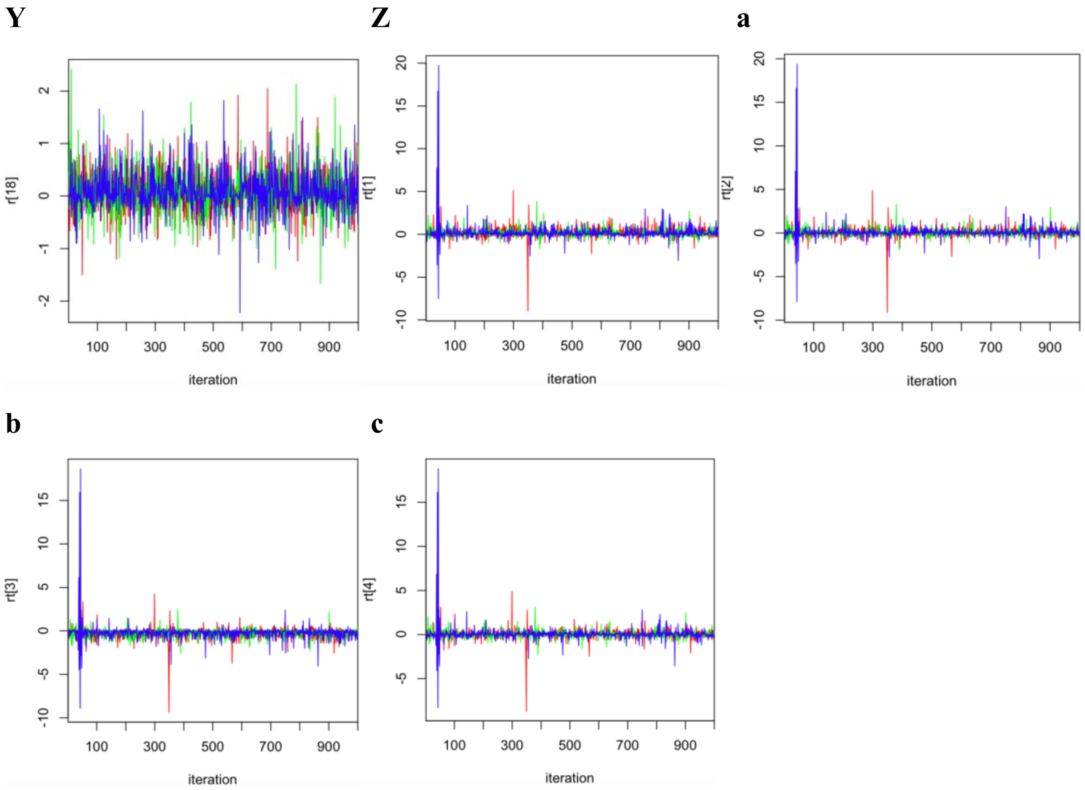



**Figure 48 Trace plot of the parameter in the spatiotemporal malaria risk modeling build by employing hierarchical Bayesian method**

Parameter distributions for (A): Intercept, (B): NDVI, (C): NDWI, (D): TWI, (E): LLIN, (F): Temperature and (G): TF, (H–Y): Location specific parameters of 18 health operational districts, (Z–c): Temporal specific parameters from 2010 to 2013
NDWI, Normalized difference water index; NDVI, Normalized difference vegetation index; TWI, Topographical wetness index; LLIN, Long lasting insecticide-treated net;

## ACRONYM LIST

| | |
|---|---|
| API | Annual Parasite Incidence |
| ABM | Agent-Based Modeling |
| CNM | National Center for Parasitology, Entomology and Malaria Control |
| CRS | Coordinate Reference System |
| CVCA | Customer Chain Value Analysis |
| DEM | Digital Elevation Model |
| DMM | Domain Mapping Matrix |
| DPHI | Department of Planning and Health Information at Ministry of Health |
| DSM | Design Structure Matrix |
| EBSMR | Standardized Morbidity Ratio Calculated by Empirical Bayes Method |
| ES | Engineering Systems |
| ES-MDM | Engineering Systems Multiple-Domain Matrix |
| FFBD | Functional Flow Block Diagram |
| GIS | Geographical Information System |
| HC | Health Center |
| HIS | Health Information System |
| HOD | Health Operational District |
| IoT | Internet of Things |
| IDW | Inverse Distance Weighed Method |
| LLIN | Long-Lasting Insecticide Treated Mosquito Net |
| LSWI | Land Surface Water Index |
| MAE | Mean Absolute Error |
| MCMC | Markov Chain Monte Carlo |
| MMP | Mobile Migrant Population |
| MoH | Ministry of Health |
| MS | Malaria Information System |
| NDVI | Normalized Difference Vegetation Index |
| NDWI | Normalized Difference Water Index |
| NGO | Non Governmental Organization |
| NP | National Program |
| PR | Parasite Rate |
| PRs | Private Sectors |
| *Pf*TSI | *Prasmodium falciparum* Temperature Suitability Index |
| RDT | Rapid Diagnostic Testing Kit |
| PHD | Provincial Health Department |
| RH | Referral Hospital |
| RMSE | Root Mean Square Error |
| SARS | Severe Acute Respiratory Syndrome |
| SMR | Standardized Morbidity Ratio |
| SoS | System of Systems |
| TWI | Topograhical Wetness Index |
| VMW | Village Malaria Worker |
| WHO | World Health Organization |