4th Intercontinental Geoinformation Days

igd.mersin.edu.tr

A correlation study for determination risk area of dengue fever and dengue hemorrhagic fever: a case study of Sisaket province, Thailand

Nutchanon Chantapoh *[1]

[1] Wuhan University, The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan, Hubei, China

Keywords
GIS
Dengue fever
Correlation

Abstract

This study has purpose on analyzing dengue fever and dengue hemorrhagic fever (DF/DHF) risk zone area in Sisaket province, Thailand, by using the subdistrict-level (Tumbon) data in sick ratio, average temperature, maximum temperature, minimum temperature, relative humidity, precipitation, population density, and housing density. The meteorological data are acquired from POWER, NASA. The data is stored in points, griding by 30 minutes of latitude and longtitude, going through the inverse distance weighting tool to interpolate the meteorological data into each Tumbon. The physical socio data are from government authority, are population from each Tumbon by monthly and housing amount from each Tumbon by yearly. Correlation analysis is used to find the correlation between sick ratio and other variables to find the risk area in the study area. Hot spots analysis (Getis ord-Gi*) is used to find the clustering area of sick ratio among each Tumbon. Results are classified into cluster maps and risk zone maps, each by yearly and monthly.

1. Introduction

Thailand had first case of DF/DHF in 1949 and a widespread epidemic in 1958 in Bangkok and Thonburi area. Since 1958, the trend of infection increases and has many forms such as every other year, 2 years then stop for 1 or every 2 years. Most of patients are 0-14 years old, with most fatal symptoms in 5-9 years old. DF/DHF cases could be found along the year, yet the highest months are in rainy season (May-August) (Insects, 2019) The effects of DF/DHF could lead to radical complications. Some patients are suffered from failure of circulatory system, shocking from leaking of plasma. Without correct medical treatment, the patient might be death in 12-24 hours. GIS helps increasing the efficiency of preventive process and cover the correct areas, deescalating the infection and death rates. The benefits from applying GIS and statistical analysis to study about the diseases can be presented to relevant agencies such as provincial public health center and hospitals.

2. Method

The patient's data from public health agency of Sisaket province, Thailand. from 2010 – 2019 The data stored in spreadsheet with headers consisting of age, gender, DF or DHF, defined date, Amphoe (district), Tumbon (subdistrict), and result of treatment (recover/dead). The data is acquired from Sisaket's public health office, ministry of public health. Population and housing data is acquired from department of provincial administration (DOPA), Ministry of Interior. Meteorological data (temperature, relative humidity, precipitation) is acquired from prediction of worldwide energy resources project (POWER), National Aeronautics and Space Administration (NASA). The used tools are ArcMap 10.5 (Hotspots Analysis (Getis-Ord Gi*), Inverse Distance Weight (IDW), Zonal Statistics as Table), JASP 0.15 (Pearson's correlation analysis), and Microsoft Excel.

* Corresponding Author

*(neononny@gmail.com) ORCID ID 0000-0002-1113-769X

Cite this study

Chantapoh, N. (2022). A correlation study for determination risk area of dengue fever and dengue hemorrhagic fever: a case study of Sisaket province, Thailand. 4th Intercontinental Geoinformation Days (IGD), 290-293, Tabriz, Iran

2.1. DF/DHF in Sisaket, 2010 – 2019.

The DF/DHF data is classified into 2 forms for separate results: yearly correlation, monthly correlation. Yearly data is the summarized number of DF/DHF incidence cases through each year in the province, and monthly data is summarized number of the disease's incidence cases through each month from studied years.

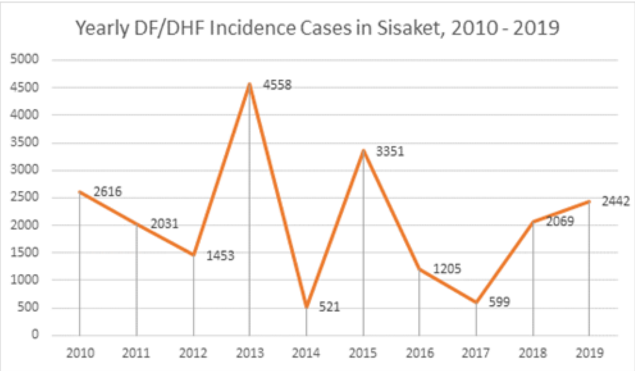

Figure 1. Yearly DF/DHF Incidence Cases in Sisaket, 2010 – 2019

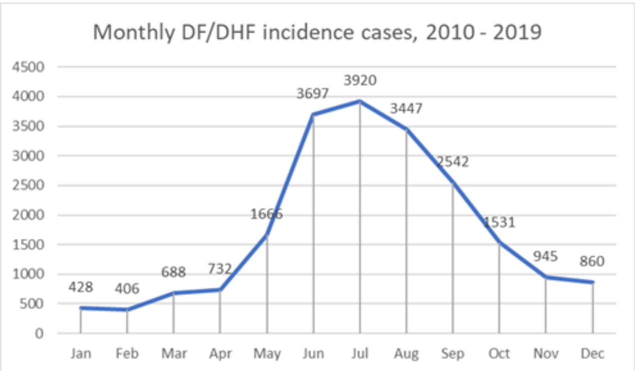

Figure 2. Monthly DF/DHF Incidence Cases in Sisaket, 2010 – 2019

2.2. Correlation Analysis

Correlation (Pearson's) tells if variables are positively related (they move in the same direction) or negatively correlated (they move in opposing directions).

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

r = r value (how much the data correlated, 1 is the highest)
x = variable 1 value (variables)
y = variable 2 value (sick ratio)

Insert variables: sick ratio, average temperature, maximum temperature (average), minimum temperature (average), relative humidity (average), precipitation amount (total/average), housing density.

2.2.1. Rating of variables' values

Rating by define the effect of each variable, 1 is for the lowest effect and 4 is for the highest effect. Find the mean value of entire data ($\bar{x}$) and its Standard Deviation (S.D.).

Table 1. Variables' values rating definition

Rating	Observed Value
4	More than $\bar{x}$ + 1 S.D.
3	Between $\bar{x}$ and less than $\bar{x}$ + 1 S.D.
2	Less than $\bar{x}$ but more than $\bar{x}$ - 1 S.D.
1	Less than $\bar{x}$ - 1 S.D.

3. Results

Analyze every variable by weighting and rating, the more affected factors will be defined with more score.

Weighting by adjust relation value (r value) out from negative value. Author used the lowest r value (in positive numeric) multiply by 2 and add to all

r value. Then, is to adjust the summarized into 1 by ratio. Spatially analyze risk area by multiply weight with rate, carried out the risk score for each Tumbon. Define class intervals by dividing the data range between maximum and minimum data with desired amount of class.

$$Class\ Interval = \frac{\text{Data}_{max} - Data_{min}}{\text{Amount of class}}$$

Table 2. Pearson's yearly correlation

Variable	r value	p value	n
Population	-0.015	0.503	2060
Avg. Temp.	0.006	0.795	
Max Temp.	0.169	< .001	
Min Temp	0.131	< .001	
Humidity	-0.229	< .001	
Precipitation	-0.11	< .001	
Housing	0.022	0.321	

Table 3. Pearson's monthly correlation

Variable	r value	p value	n
Population	-0.015	0.45	2472
Avg. Temp.	-0.088	< .001	
Max Temp.	-0.416	< .001	
Min Temp.	0.515	< .001	
Humidity	0.561	< .001	
Precipitation	0.658	< .001	

Results are fully displayed in risk zone maps by years 2010 – 2019 and by months from studied years. The example figures (figure 3 and figure 4) show the comparison between the years and months with the highest sick ratio with the lowest.

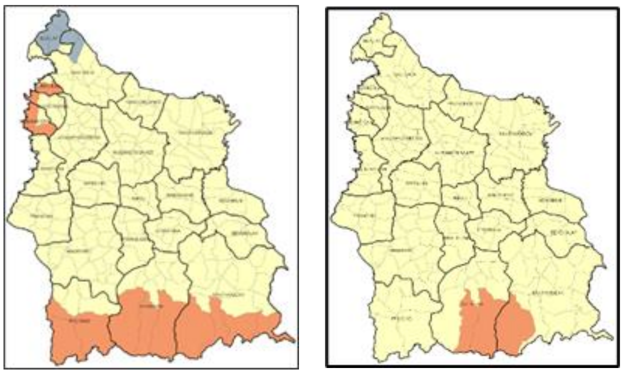

Figure 3. (From left to right) Risk zone maps of 2013, 2014

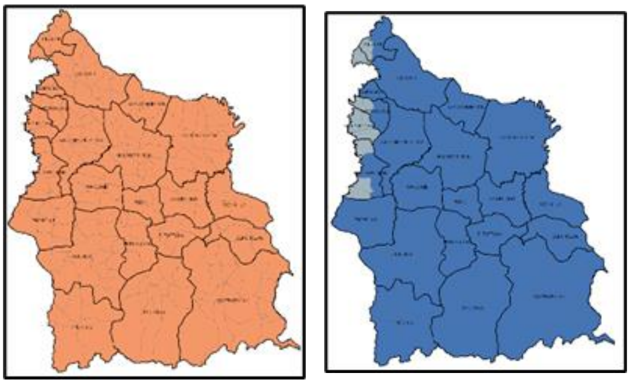

Figure 4. (From left to right) Risk zone maps of July, January

Table 4. Correlation between yearly sick ratio and risk score

Year	Sick Ratio	Average Risk Score	r score	p value
2010	179.82	2.6210	0.213	0.554
2011	139.86	1.7005		
2012	99.63	2.7474		
2013	311.76	2.5663		
2014	35.56	2.4568		
2015	228.15	2.6981		
2016	81.95	2.6843		
2017	40.69	2.1679		
2018	140.46	2.1174		
2019	165.8	3.3073		

Table 5. Correlation between monthly sick ratio and risk score

Month	Sick Ratio	Average Risk Score	r score	p value
Jan	28.2	1.5508	0.799	0.002
Feb	27.98	1.5076		
Mar	46.97	2.1720		
Apr	50.14	2.4219		
May	112.64	3.3879		
Jun	252.43	3.1727		
Jul	266.21	3.0612		
Aug	231.39	3.2425		
Sep	168.43	3.2425		
Oct	100.39	2.1813		
Nov	61.29	2.1204		
Dec	53.96	1.3356		

4. Discussion

The study of correlation and estimate risk area of DF/DHF by gathering data from agencies, has purpose for find relationship between factors and produce risk zone maps of Sisaket province. The studied data is 10-year long (2010 – 2019), is the DF/DHF incidence cases, had defined the factor that might cause the diseases. There are 6 variables involve in the study which are population density, average temperature, average maximum temperature, average minimum temperature, precipitation amount, and housing density (in monthly analysis used average precipitation instead of total amount and have no monthly housing density). For Tumbon (subdistrict) level, the meteorological data might not have wide range and much difference. The estimation from geoinformatics tools is not exactly accurate, unless the data are physically acquired from local area.

For statistical analysis results, yearly analyzing gave average maximum and minimum temperature 16.9% and 13.1% significantly positive correlate ($p<0.001$), as well as average relative humidity and precipitation 22.9% and 11% significantly negative correlate. Meanwhile, monthly analyzing gives average temperature and maximum temperature significantly negative correlate with 8.8% and 41.6%, while average minimum temperature, relative humidity and precipitation are significantly positive correlate with 51.5%, 56.1%, and 65.8%. ($p<0.001$). The results obviously show that monthly study provide stronger statistical correlations more than annually study. Moreover, the average risk scores show that the monthly study also provide more correlation than yearly study (79.9% and 21.3%, monthly and yearly).

5. Conclusion

The yearly correlation results in positive with maximum and minimum temperature and negative with relative humidity and precipitation, which all factor's trends are rising each year. Positive correlation r values are 16.9% and 13.1% and negative correlation r values are 22.9% and 11%. This phenomenon could lead to decrease of DF/DHF sick rate.

The monthly correlation show that maximum temperature and minimum temperature are respectively negative and positive correlated. This could lead to find the suitable temperature range for DF/DHF incidence. Moreover, the relative humidity and precipitation are both strongly positive correlated, summarizing that seasonal meteorological data are suited for study the correlation with the diseases. The most affected variables are precipitation, relative humidity, minimum temperature, and maximum temperature (negative correlated) respectively. These factors are more than 40% correlated with the sick ratio.

The monthly risk zone maps (which provide significant correlation with sick ratio) show that through the year, DF/DHF are riskier in the southern part of the studied area, Sisaket province. PHU SING, KHUN HAN, and KANTHARALAK Amphoes are the 3 Amphoes located in the south of the province, northern to the Thai-Cambodian border with a vast area of mountain forest. The study of correlation in monthly summarization provides more correlative results than yearly study. The office of provincial public health should consider in providing more density measurements in prevention and healthcare facilities.

Acknowledgement

I would like to express my appreciation to Dr. Phattraporn Soytong, professor of Burapha University and Prof. SHU Hong, professor at Wuhan university, mr. Thanaphat Thanaratkiettikun and Dr. Pramet Kaewmesri from Geo-Informatics and Space Technology Development Agency (GISTDA), Mr. Pichet Peudkhuntod, and Ms. Kanyarat Lhaosatienkit from Sisaket public health office, Dr. Wittawin Chantapoh, Ms. Varittha

Soposuk, Royal Thai survey department (RTSD), and Sirindhorn Center of Geoinformatics (SCGI) master program batch 3 colleagues for all assistance given to me.

References

Atiwath Waraputh. (2010). Factors Associated with Prevention and Control Behavior Toward Dengue Hemorrhagic Fever Among People Srisaket Province. (Master of Public Health Promotion). Ubon Ratchathani Rajabhat University

Ausawadee Thawara, Apiwat Tawatsin, Jakkrawarn Sirisopa, & Phayu Phakdeenuan. (2016). Mosquitoes. Ministry of Public Health

Budsara Thepawattanasuk. (2013). Dengue Fever and Dengue Hemorrhagic Fever in Chaoprayayomraj Hospital. Journal of Bamrasnaradura Infectious Diseases Institute, 7.

Bureau of Communicable Diseases Led by Insects. (2019). Dengue Fever Situation Report. Department of Disease Control, : Ministry of Public Health,

Jiratiwan Kruasilp. (2006). Geographic Information System for Determining Dengue Fever Risk Areas in Changwat Nakhon Ratchasima. (Master of Arts Program in Geography). Chulalongkorn University,

Keerati Ponpetch. (2008). Geographic Information System for Prevention and Control Planning of Dengue Haemorrhagic Fever in Phetchaburi Province. (Master of Science). Silpakorn University, Graduate School.

Moustafa A. Hegazi, Marwan A. Bakarman, Turki S. Alahmadi, Nadeem S. Butt, Ahmed M. Alqahtani, Badr S. Aljedaani, & Abdulrahman H. Almajnuni. (2020). Risk Factors and Predictors of Severe Dengue in Saudi Population in Jeddah, Western Saudi Arabia: A Retrospective Study. The American Society of Tropical Medicine and Hygiene.

Muhammad Shahzad Sarfraz, Nitin K Tripathi, Taravudh Tipdecho, Thawisak Thongbu, Pornsuk Kerdthong, & Souris, M. (2012). Analyzing the spatio-temporal relationship between dengue vector larval density and land-use using factor analysis and spatial ring mapping. Bio Med Central.

Natharin Dechasiripong. (2014). Mathematical Model for Analysis of Dengue Fever Spreading Risk Areas Based on Climatic Factors. (Doctor of Philosophy in Social Development and Environmental Management). National Institute of Development Administration,

Prangthip Buafueang. (2013). A study of the application of Geographic Information System (GIS) for studying the spread of Dengue Haemorrhagic Fever (DHF) outbreak within Tamai District, Chanthaburi. (Master of Science (Environmental Management). National Institute of Development Administration,

Teeranai Srithamarong. (2016). Crime Risk Area Assesment Based Geo-Spatial Statistics. (Master of Science Program in Environmental Science Thesis). Silpakorn University, Graduate School.