

Hotspot Analysis of Dengue Hemorrhagic Fever in Ho Chi Minh City, Vietnam

Thi-Tuyet-Mai Nguyen¹, Huong-Giang Nguyen¹, Thi-Hien Cao²

¹Faculty of Pharmacy, East Asia University of Technology, Hanoi, Vietnam

²Faculty of Nursing, East Asia University of Technology, Hanoi, Vietnam

Corresponding Author: Thi-Tuyet-Mai Nguyen

DOI: <https://doi.org/10.52403/ijshr.20230358>

ABSTRACT

Background: Dengue Hemorrhagic Fever (DHF) in Vietnam remains a serious emerging arboviral disease, which generates significant concerns among international health authorities. Incidence rates of DHF have increased significantly during the last few years in many provinces and cities, especially Ho Chi Minh. The purpose of this study is to carry out a study on analysis of hotspots of DHF in the Ho Chi Minh city.

Methods: The Getis-Ord's G_i^* statistic-based hotspot analysis was employed to study spatial patterns (hotspots and coldspot) of DHF incidence.

Results: a case study from a dataset of DHF collected in Ho Chi Minh city has indicated that a total of four hotspots of DHF were successfully detected. These hotspots were identified in Hoc Mon, District 12, Binh Tan, and District 7. In addition, 18 districts were not statistically significant at the level of 0.05.

Conclusions: There was evidence for the existence of statistically significant clusters of DHF in Ho Chi Minh. The study results prove the effectiveness of Getis-Ord's G_i^* statistic-based hotspot analysis in studies of DHF. Findings in this study serves as a starting point for further research into the social and environmental factors influencing spatial patterns of DHF.

Keywords: Hotspot analysis and Coldspot analysis, Dengue Hemorrhagic Fever, Getis-Ord's G_i^* statistic, Ho Chi Minh, Vietnam.

INTRODUCTION

Dengue virus (DENV, family Flaviviridae), which recently experienced its greatest

DENV outbreak in over ten years in Vietnam's largest city, Ho Chi Minh City, is a serious threat to the world's public health. It causes significant morbidity and mortality in hyperendemic nations like Vietnam (1). With more than 2.5 billion inhabitants living in areas at high risk of infection from the dengue virus, it is one of the most significant vector-borne viruses in the world. The majority of those at risk are from inter-tropical nations like Vietnam (2). Four genetically and antigenically diverse serotypes of DENV, known as DENV-1 through DENV-4, may be distinguished from one another. These various genotypes have diversified into different populations with variable levels of geographic persistence (3). DENV is thought to be the cause of 105 million cases worldwide each year, with most cases being asymptomatic (4). DENV results in serious economic harm (5). Symptoms of DENV infection can range from mild, self-limited acute febrile illness to severe organ dysfunction and shock (6). It is therefore, studies of DHF hotspots play an important role in DHF control.

Dengue Hemorrhagic Fever (DHF) in Vietnam remains a serious emerging arboviral disease. With an estimated 1.6 million infections per year, DHF is widespread throughout Vietnam in urban and periurban areas (5). As is typical in hyperendemic countries, prior research has demonstrated that all four DENV serotypes had circulated in Vietnam at some point, with DENV-1 and DENV-2 being the most often

found serotypes (7). Up until 2018, DENV-1 was the predominant serotype in Ho Chi Minh City and the nearby areas (1). A large DENV outbreak occurred in Ho Chi Minh city in 2022, with 78,561 dengue cases recorded there by the Ho Chi Minh Centre for Disease Control between January 1st to December 11th (8). Numerous reasons, such as environmental factors, decreased population immunity, the introduction of novel serotypes or genotypes, as well as the potential diversification of currently circulating lineages, may be linked to this huge outbreak (1). The commonly used methods for the detection of hotspots were based on local Moran's I, followed by *Getis-Ord's* G_i^* statistic, Kulldorff's spatial scan statistic and Kernel density (9). Among these statistics, *Getis-Ord's* G_i^* statistic is one of the most commonly used for hotspot analysis. For example, using *Getis-Ord's* G_i^* statistic and geographically weighted principal component analysis, the impact of living environment deprivation on COVID-19 hotspot was successfully examined in Kolkata megacity, India (10). The local *Getis-Ord's* G_i^* statistic was successfully applied to detect hotspots and coldspots of COVID-19 cases in four waves in Vietnam (11,12). In the United States, the combination of *Getis-Ord* G^* statistic and artificial neural network in a GIS environment was employed by (13) to model COVID-19 incidence rates across the Continental United States. With the aims of analyzing the geographical distribution of COVID-19 and of identifying high-risk areas for the occurrence of cases and deaths from the disease in the indigenous population of Brazil, both of *Getis-Ord* *General G* and *Getis-Ord* G^* statistics were successfully used by Alves et al. (2021) to identify high spatial risk clusters from 32,041 cases of and 471 deaths by COVID-19 (14). In Italy, widely used spatial indicators such as the local *Moran's I*, *Getis and Ord G* statistics, *LISA* clusters, and *Moran* scatterplots were also successfully used by Ghosh and Cartone (2020) to identify spatial effects and spatio-

temporal patterns of the outbreak of COVID-19 from February to June 2020 (15). Most recently, with the help of geostatistical techniques in spatial prediction and mapping of the COVID-19 hotspots in India, four factors (total population, population density, foreign tourist arrivals, and COVID-19 confirmed cases) were considered in the detection of COVID-19 hotspots (16). However, there is currently a lack of studies on DHF hotspots and its dynamics in Ho Chi Minh. As such, this study focuses on the hotspots of DHF incidence, particularly the high number of DHF cases reported in Ho Chi Minh city.

The objective of this study is to analyze DHF hotspots in Ho Chi Minh city, Vietnam. A datasets of DHF cases were first collected. Spatial weight matrix was then constructed using the first order order of contiguity. The *Getis-Ord's* G_i^* statistic-based hotspot analysis was finally employed to detect hotspots of DHF incidence in the study area.

MATERIALS & METHODS

Materials

Data used in the study were collected from Ho Chi Minh Center for Disease Control. The dataset is the cumulative number of DHF cases collected in Ho Chi Minh city. The spatial distribution of DHF is shown in figure 1 (right). Data from figure 1 (right) illustrates that a large number of infections are mainly concentrated in the city centre with a range of from 654 to 1032 cases. The areas with low number of DHF infections were mainly concentrated in the south of the city including 5 districts with a lower number of infections of less than 274 cases. Data on from Figure 1 (right) shows the spatial distribution of DHF incidence (the number of DHF infections per 100,000 inhabitants). Data in Figure 1 (right) also shows that the southwest districts of the city had a higher DHF infection rate than other areas. The areas with low DHF infection rates were mainly located in the southern and central, northern and eastern parts of the city, respectively.

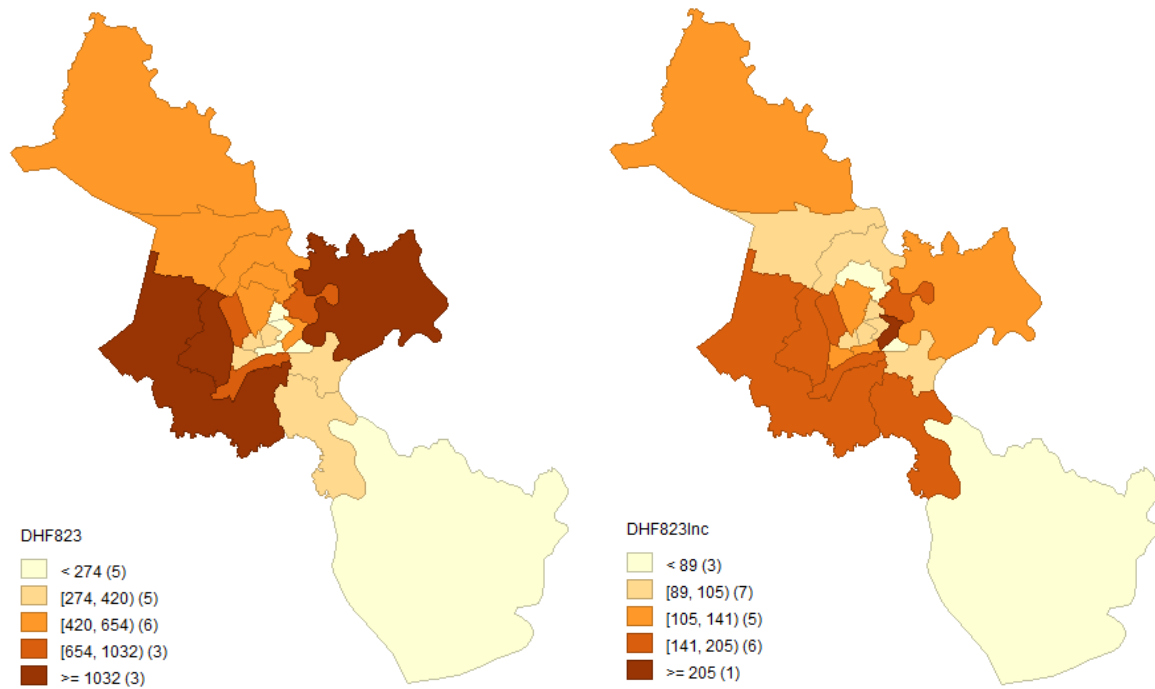


Figure 1. Spatial distribution of DHF cases (left) and incidence (right) in Ho Chi Minh city.

Methods

There are different methods for analyzing spatial patterns and detecting hotspots including spatial autocorrelation and cluster analysis (17). A hotspot can be defined as an area that has higher concentration of events compared to the expected number given a random distribution of events (18). The application of hotspot analysis within public health and epidemiological research as well as in other disciplines (e.g., a great deal of the literature on hotspot analysis comes from crime mapping and research) has increased significantly in the past couple of decades mainly due to the advent of Geographic Information Systems (GIS)-based software (19–21).

Spatial autocorrelation analysis looks at how similar are those values that are closer to each other (22). Measures of spatial autocorrelation can be categorized as global or local indicators of spatial association (23). Spatial autocorrelation analysis was performed on the incidence rates to test whether the cases were distributed randomly over space and, if not, to evaluate any identified spatial disease clusters for statistical significance (24). A hotspot is defined as a condition indicating some form of clustering in a spatial distribution (24).

Hotspot analysis is based on the *Getis-Ord's* G_i^* statistic. Hotspot analysis characterizes the presence of hotspots (high clustered values) and coldspots (low clustered values) over an entire area by looking at each feature within the context of its neighboring features (18). Hotspot can separate clusters of high values from cluster of low values. It is, therefore, *Getis-Ord's* G_i^* statistic was used to identify the counties of high and low numbers of COVID-19 cases (11,12). The form of *Getis-Ord's* G_i^* statistic is defined as follows (25):

$$G_i^* = \frac{\sum_{j=1}^N W_{ij} x_j - \bar{x} \sum_{j=1}^N W_{ij}}{S \sqrt{\frac{N \sum_{j=1}^N [W_{ij}^2 - (W_{ij})^2]}{N - 1}}} \quad (1)$$

with:

$$\bar{x} = \frac{1}{N} \sum_{j=1}^N x_j \quad (2)$$

and:

$$S = \sqrt{\frac{\sum_{j=1}^N x_j^2}{N} - (\bar{x})^2} \quad (3)$$

Expectation:

$$E(G_i^*) = \frac{W_i^*}{n-1} \quad (4)$$

with:

$$W_i^* = \sum_{j=1}^n w_{ij}(d) \quad (5)$$

and variance:

$$\text{Var}(G_i^*) = \frac{W_i^*(n - W_i^*)Y_{i2}^*}{n^2(n-1)(Y_{i1}^*)^2} \quad (6)$$

with:

$$Y_{i1}^* = \frac{\sum_{j=1}^n x_j}{n}; \quad (7)$$

and:

$$Y_{i2}^* = \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i x_j)^2}{n} - (Y_{i1}^*)^2; \quad (8)$$

where: the *Getis-Ord's* G_i^* statistic is computed for the number of COVID-19 cases at county i ; x_i , x_j , \bar{x} , and W_{ij} are defined in equation (1); and N is the total number of neighborhood counties as defined in equation (2). W_{ij} can be constructed using the methods of the first order and second of contiguity. In this study, adjacency to compute W_{ij} is defined using the first order of contiguity.

The *Getis-Ord's* G_i^* coefficient at county i (G_i^*) also ranges between -1 and +1. If $G_i^* > 0$ and $p(G_i^*) < \alpha$ then there exists a spatial clustering of high-high values (11,12). In this case, these high-high values, so-called a hotspots, reflects the presence of high numbers of COVID-19 cases among county i and its neighborhood counties ($j \in J_i$). Whereas, if $G_i^* < 0$ and $p(G_i^*) < \alpha$ then there exists a spatial clustering of low-low values (11,12). These low-low values are called a coldspots indicating low numbers of COVID-19 cases among county i and its neighborhood counties ($j \in J_i$). Similar to those in the definition of local Moran's I statistic, if the value of G_i^* close zero and $p(G_i^*) < \alpha$ then there will be neither hotspots nor coldspots or random distribution of COVID-19 cases (11). Clusters of cases that occur randomly can also have an influence on the spread of an infectious disease (24). The output from *Getis-Ord's* G_i^* statistic identifies spatial clusters of high values

(hotspots) and spatial clusters of low values (cold spots).

The high/low clustering tool is an inferential statistic, which means that the results of the analysis are interpreted within the context of a null hypothesis (24). A lots of attempts have been put on the use of *Getis-Ord's* G_i^* statistic with the help of ArcGIS software using *Getis z-scores* (14,18,26) defined in a study by Mitchel (26). However, as discussed above, the presence of a strongly skewed distribution in the dataset fails the test. It is, therefore, testing for the significance of the *Getis-Ord's* G_i^* statistic in this study was also carried out by a randomization test using 999 permutations. In this work, with the help of the spatial statistics software, GeoDA, developed by (27), a randomization test was used to test the significance of spatial autocorrelation statistics. Spatial autocorrelation statistics were generated and tested at the significance of 0.05 using 999 permutations. When the p-value is very small, it means it is very unlikely (small probability) that the observed spatial pattern is the result of random processes, so the null hypothesis can be rejected (24).

RESULTS AND DISCUSSIONS

Analysis of the properties of spatial weight matrix

In this study, the first order of contiguity was used to determine the spatial weight matrix. The attribute values of the spatial weight matrix including the connection map, connection graph and number of neighbors are shown in Figures 2 and 3. In this study, a total of 22 districts in Ho Chi Minh City were involved in the computation. The minimum, mean, median, and maximum numbers of neighbors were 1, 4.91, 5.0 and 7, respectively. Whereas, the percentage of districts with non-zero was 22.31. Data in Figure 3 shows that several districts have many neighbors, such as 4 districts with 7 neighbors, 10 districts with 5 neighbors. On the other hands, there are some districts with a low number of neighbors, such as 2 districts with only one neighbor and one district with three neighbors.

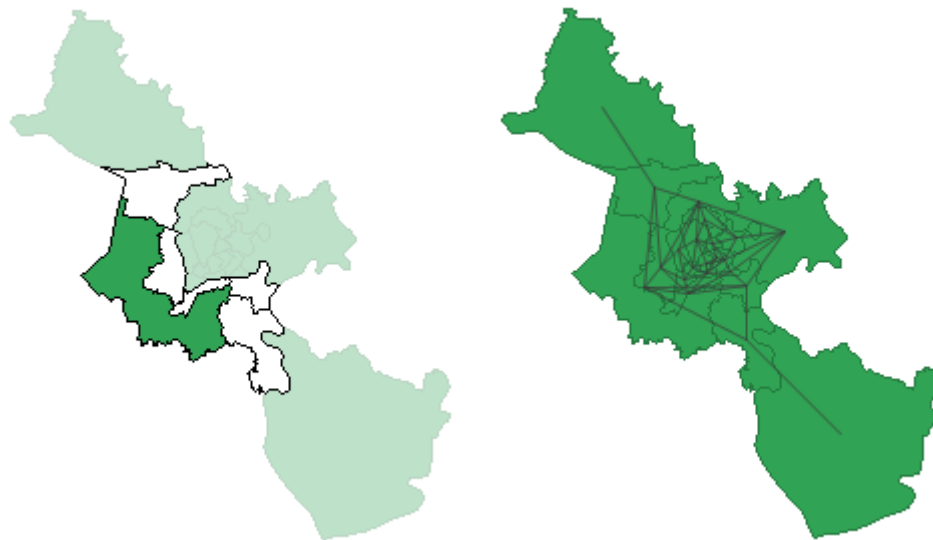


Figure 2. Connectivity map (left) and connectivity graph (right) in Ho Chi Minh city.

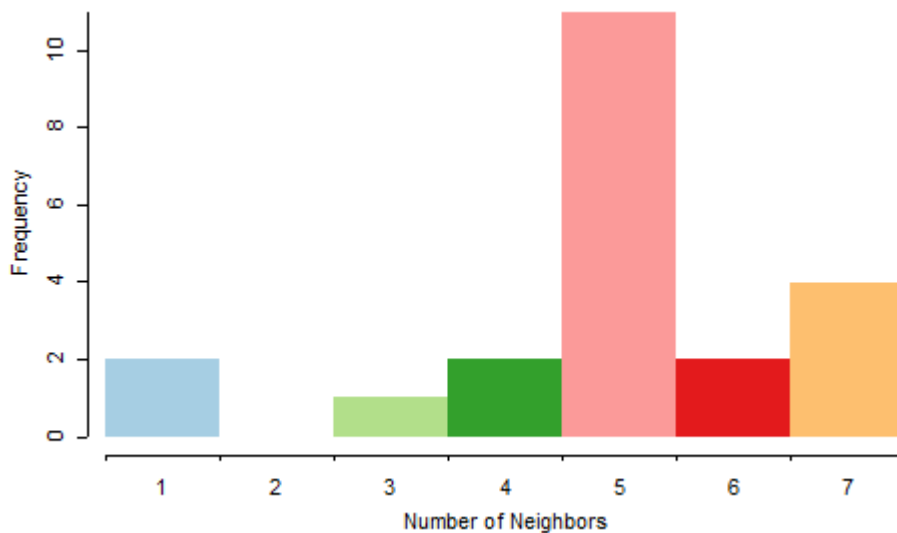


Figure 3. Histogram of neighbors for districts in Ho Chi Minh city.

Hotspot analysis of dengue hemorrhagic fever

Hotspots of DHF incidence in the case of using first order of contiguity were shown in Figure 4. Data from the cluster map in Figure 4 (left) shows that a total of 4 hotspots were detected. A total of 18 districts were not statistically significant at the level of 0.05. Hotspots were mainly distributed in urban districts in the centre of Ho Chi Minh City with their corresponding DHF infection rates as follows: Hoc Mon (97 cases per 100,000 inhabitants), District 12 (102 cases per 100,000 inhabitants), Binh Tan (147 cases per 100,000 inhabitants), and District 7 (99 cases per 100,000 inhabitants). It is worth noting that some districts have high DHF

infection rates but no hotspots of DHF were detected such as District 1 (205 cases per 100,000 inhabitants), Binh Chanh and Nha Be (162 cases per 100,000 inhabitants) or District 8 (161 cases per 100,000 inhabitants). The reason for this is that their surrounding districts had a smaller number of DHF cases, so the *Getis-Ord's* G_i^* does not consider them as DHF hotspots. In addition, no DHF coldspots were also detected although there were some districts having low DHF infection rates such as Go Vap (74 cases per 100,000 inhabitants), Can Gio (74 cases per 100,000 inhabitants), District 4 (79 cases per 100,000 inhabitants) and District 11 (89 cases per 100,000 inhabitants). Besides that. Data from Figure 3 (left)

demonstrate that 18 districts were unstatistically significant at the level of 0.05. These districts are mainly distributed in the north and south of the city. Data from Figure 3 (right) also shows the four levels of statistical significance for districts of Ho Chi Minh city. Levels of statistical significance are expressed at four different scales: statistical unsignificance (p-value > 0.05)

and statistical significance at levels of 0.05, 0.01 and 0.001. Data from Figure 3 (right) illustrate that District 7 was the only one statistically significant at a very high level (p-value < 0.001). One district, namely District 12, was statistically significant at a high level (p-value < 0.01). Two districts, Hoc Mon and Binh Tan, were statistically significant at the level of 0.05.

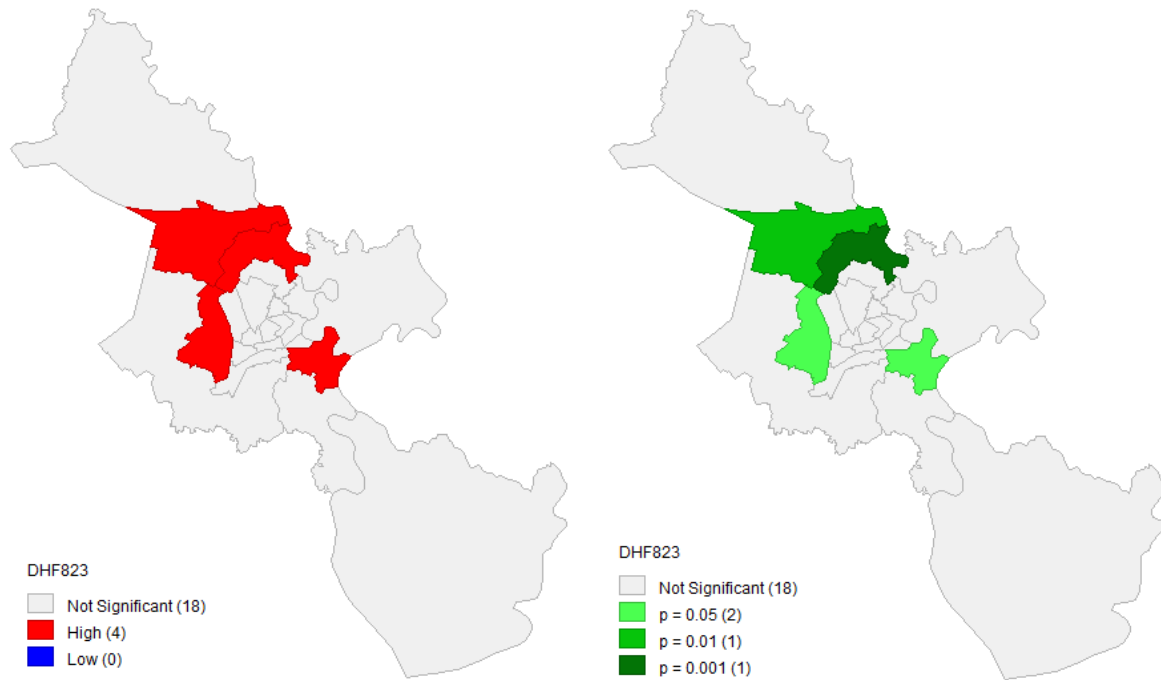


Figure 4. DHF cluster map (left) and significance map (right) in Ho Chi Minh city.

CONCLUSION

With the help of Geographic Information System, this study aimed to perform the analysis of hotspots of DHF in the Ho Chi Minh city. The *Getis-Ord's G_i^{*}* statistic-based hotspot analysis was employed to measure spatial patterns of DHF incidence including hotspots and coldspot. It was found a case study from a datasets of DHF incidence in Ho Chi Minh city, four hotspots of DHF were successfully detected in Hoc Mon, District 12, Binh Tan, and District 7. Moreover, 18 districts were not statistically significant at the level of 0.05. There was evidence for the existence of statistically significant clusters of DHF in Ho Chi Minh city. The study results prove the effectiveness of *Getis-Ord's G_i^{*}* statistic-based hotspot analysis in studies of DHF clusters. Findings in this study serves as a

starting point for further research into the social and environmental factors influencing spatial patterns of DHF. Since hotspot analysis can help better allocate resources to combat the DHF outbreak, it should be widely employed in DHF surveillance. This endeavor will support the dengue prevention plan. Finally, we anticipate that this research will serve as a paradigm for future spatial models that are more accurate. And in order to prevent and forecast dengue spread over an expansion area, public health officials might use the DHF cluster map that was obtained from the study. The cluster map might also be useful for public alerts and awareness.

Declaration by Authors

Acknowledgement: The authors would like to thank Ho Chi Minh Center for Disease

Control (Ho Chi Minh CDC) for providing the data.

Source of Funding: None

Conflict of Interest: None

REFERENCES

1. Tran VT, Inward RPD, Gutierrez B, Nguyen NM, Rajendiran I, Thanh PN, et al. Cryptic transmission and re-emergence of Cosmopolitan genotype of Dengue Virus Serotype 2 within Ho Chi Minh City and Southern Vietnam. medRxiv. 2023;2004–23.
2. Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, et al. Dengue: a continuing global threat. Nat Rev Microbiol. 2010;8(Suppl 12):S7–16.
3. Messina JP, Brady OJ, Scott TW, Zou C, Pigott DM, Duda KA, et al. Global spread of dengue virus types: mapping the 70 year history. Trends Microbiol. 2014;22(3):138–46.
4. Chatchen S, Sabchareon A, Sirivichayakul C. Serodiagnosis of asymptomatic dengue infection. Asian Pac J Trop Med. 2017;10(1):11–4.
5. Cattarino L, Rodriguez-Barraquer I, Imai N, Cummings DAT, Ferguson NM. Mapping global variation in dengue transmission intensity. Sci Transl Med. 2020;12(528):eaax4144.
6. Yacoub S, Wills B. Dengue: an update for clinicians working in non-endemic areas. Clin Med (Northfield Il). 2015;15(1):82.
7. Ty Hang VT, Holmes EC, Veasna D, Quy NT, Tinh Hien T, Quail M, et al. Emergence of the Asian 1 genotype of dengue virus serotype 2 in viet nam: in vivo fitness advantage and lineage replacement in South-East Asia. PLoS Negl Trop Dis. 2010;4(7):e757.
8. HCDC. Hand, foot, mouth and dengue fever epidemic situation updated until September 11, 2022 [Internet]. 2022. Available from: <https://hcdc.vn/tinh-hinh-dich-benh-tay-chan-mieng-sot-xuat-huyet-cap-nhat-den-ngay-1192022-81472dae343782d142f3940ac20491df.html>
9. Nazia N, Butt ZA, Bedard ML, Tang W-C, Sehar H, Law J. Methods used in the spatial and spatiotemporal analysis of COVID-19 epidemiology: a systematic review. Int J Environ Res Public Health. 2022;19(14):8267.
10. Das A, Ghosh S, Das K, Basu T, Dutta I, Das M. Living environment matters: Unravelling the spatial clustering of COVID-19 hotspots in Kolkata megacity, India. Sustain Cities Soc. 2021;65:102577.
11. Vu D-T, Nguyen T-T, Hoang A-H. Spatial clustering analysis of the COVID-19 pandemic: A case study of the fourth wave in Vietnam. Geogr Environ Sustain. 2021;14(4).
12. Hoang A, Nguyen T. Identifying Spatio-Temporal Clustering of the COVID-19 Patterns Using Spatial Statistics: Case Studies of Four Waves in Vietnam. Int J Appl Geospatial Res. 2022;13(1):1–15.
13. Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. Sci Total Environ. 2020;728:138884.
14. Alves JD, Abade AS, Peres WP, Borges JE, Santos SM, Scholze AR. Impact of COVID-19 on the indigenous population of Brazil: A geo-epidemiological study. Epidemiol Infect. 2021;149:e185.
15. Ghosh P, Cartone A. A Spatio-temporal analysis of COVID-19 outbreak in Italy. Reg Sci Policy Pract. 2020;12(6):1047–62.
16. Parvin F, Ali SA, Hashmi SNI, Khatoun A. Accessibility and site suitability for healthcare services using GIS-based hybrid decision-making approach: a study in Murshidabad, India. Spat Inf Res. 2021;29:1–18.
17. Elliott P, Wartenberg D, Marshall RJ, Rosli NM, Shah SA, Mahmood MI, et al. A review of methods for the statistical analysis of spatial patterns of disease. J R Stat Soc Ser A (Statistics Soc. 2021;154(1):998–1006.
18. Nguyen TT, Vu TD. Use of hot spot analysis to detect underground coal fires from landsat-8 TIRS data: A case study in the Khanh Hoa coal field, North-East of Vietnam. Environ Nat Resour J. 2019;17(3).
19. Murad A, Khashoggi BF. Using GIS for disease mapping and clustering in Jeddah, Saudi Arabia. ISPRS Int J Geo-Information. 2020;9(5):328.
20. Gao S, Mioc D, Anton F, Yi X, Coleman DJ. Online GIS services for mapping and sharing disease information. Int J Health Geogr. 2008;7:1–12.
21. Rican S, Salem G. Mapping disease. A companion to Heal Med Geogr. 2009;14:96.
22. Vu D-T, Nguyen T-T. Spatial pattern of land surface temperatures and its relation to

- underground coal fires in the Khanh Hoa Coal Field, North-East of Vietnam. Arab J Geosci. 2021;14(3).
23. Fecht D, Hansell AL, Morley D, Dajnak D, Vienneau D, Beevers S, et al. Spatial and temporal associations of road traffic noise and air pollution in London: Implications for epidemiological studies. Environ Int. 2016;88:235–42.
24. Kulldorff M, Feuer EJ, Miller BA, Freedma LS. Breast cancer clusters in the northeast United States: a geographic analysis. Am J Epidemiol. 1997;146(2):161–70.
25. Cliff AD, Ord JK. Spatial processes: models & applications. (No Title). 1981;
26. Mitchell A. The ESRI guide to GIS analysis: geographic patterns & relationships. Vol. 1. ESRI, Inc.; 1999.
27. Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. In: Handbook of applied spatial analysis: Software tools, methods and applications. Springer; 2009. p. 73–89.
- How to cite this article: Thi-Tuyet-Mai Nguyen, Huong-Giang Nguyen, Thi-Hien Cao. Hotspot analysis of dengue hemorrhagic fever in Ho Chi Minh City, Vietnam. International Journal of Science & Healthcare Research. 2023; 8(3): 436-443. DOI: [10.52403/ijshr.20230358](https://doi.org/10.52403/ijshr.20230358)
