Science of the Total Environment 598 (2017) 1151–1159

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Ascertaining the impact of public rapid transit system on spread of dengue in urban settings

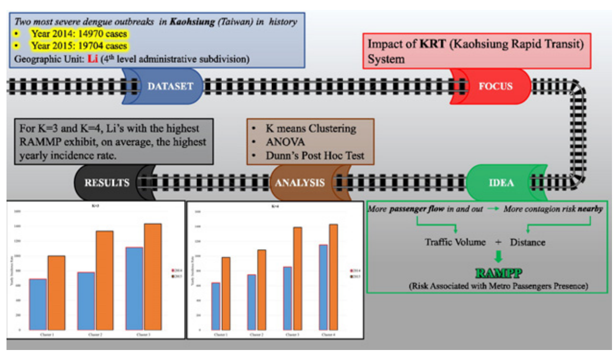
Mattia Sanna [a], Ying-Hen Hsieh [a,b,*]

[a] Department of Public Health, China Medical University, Taichung 40402, Taiwan
[b] Center for Infectious Disease Education and Research, China Medical University, Taichung 40402, Taiwan

HIGHLIGHTS

- Combined effect of proximity and metro passenger flow on dengue spread is studied.
- Incidence is significantly associated with station proximity and passenger volume.
- Large passenger flow of station could increase contagion risk in the neighborhood.
- Result could be useful in future design of effective intervention/control measures.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 11 November 2016
Received in revised form 21 March 2017
Accepted 7 April 2017
Available online 8 May 2017

Editor: D. Barcelo

Keywords:
Dengue
Population density
Mobility
Transport system
Urban
Taiwan

ABSTRACT

Urbanization is an important factor contributing to the global spread of dengue in recent decades, especially in tropical regions. However, the impact of public transportation system on local spread of dengue in urban settings remains poorly understood, due to the difficulty in collecting relevant locality, transportation and disease incidence data with sufficient detail, and in suitably quantifying the combined effect of proximity and passenger flow. We quantify proximity and passenger traffic data relating to 2014–2015 dengue outbreaks in Kaohsiung, Taiwan by introducing a "Risk Associated with Metro Passengers Presence" (RAMPP), which considers the passenger traffic of stations located within a fixed radius, giving more weight to the busier and/or closer stations. In order to analyze the contagion risk associated with nearby presence of one or more Kaohsiung Rapid Transit (KRT) stations, we cluster the Li's (the fourth level administrative subdivision in Taiwan) of Kaohsiung based on their RAMPP value using the K-means algorithm. We then perform analysis of variance on distinct clusterings and detect significant differences for both years. The subsequent post hoc tests (Dunn) show that yearly incidence rate observed in the areas with highest RAMPP values is always significantly greater than that recorded with smaller RAMPP values. RAMPP takes into account of population mobility in urban settings via the use of passenger traffic information of urban transportation system, that captures the simple but important idea that large amount of passenger flow in and out of a station can dramatically increase the contagion risk of dengue in the neighborhood. Our study provides a new perspective in identifying high-risk areas for transmissions and thus enhances our understanding of how public rapid transit system contributes to disease spread in densely populated urban areas, which could be useful in the design of more effective and timely intervention and control measures for future outbreaks.

* Corresponding author at: Department of Public Health, China Medical University Taichung, 91 Hsueh-Shih Road, Taichung 40402, Taiwan.
E-mail address: hsieh@mail.cmu.edu.tw (Y.-H. Hsieh).

http://dx.doi.org/10.1016/j.scitotenv.2017.04.050

1. Introduction

Dengue fever is a vector-borne infectious disease mainly spread by mosquitoes *Aedes aegypti*, and *Aedes albopictus* to a lesser degree. Dengue is currently considered an international public health emergency (WHO and TDR, 2009) because of its extremely fast diffusion rate. According to a recent estimate, around 390 million people are infected annually with dengue virus worldwide (Bhatt et al., 2013). Dengue is considered the most important arboviral disease of humans, with estimated over half of the world's population living in areas of contagion risk (Gubler, 2011). The frequency and magnitude of dengue epidemic have increased dramatically during the past 40 years, during which time both the dengue virus and the mosquito vectors have expanded geographically in most tropical and even some subtropical regions of the world. Modern urbanization in the recent decades, especially in the tropical and subtropical regions, stood out as one of the principal drivers of the epidemic dengue, along with globalization and lack of effective mosquito control (Gubler, 2011).

The abundance of reported cases is often not sufficient to quantify the risk of dengue infection, neglecting for instance the impact of human behavior (Wen et al., 2015). Moreover, in addition to environmental factors, the extensive spread of the disease might be explainable by other factors (Wen et al., 2012; Hsieh et al., 2013). The importance of human mobility affecting infectious disease spread and emergence is widely recognized (Prothero, 1977; Wilson, 1995; Colizza et al., 2007; Hsieh et al., 2007; Khan et al., 2009). In particular, the impact of public transportation network on local spread of infectious diseases, especially that of Tuberculosis, has been extensively studied (Tatem et al., 2006; Xu et al., 2013; Horna-Campos et al., 2007; Barrett et al., 2008; Edelson and Phypers, 2011). Recent studies suggest that the routine movement of individuals could potentially enhance dengue transmission, as it can easily go beyond the dispersal range of mosquito population, resulting in large-scale outbreaks (Wen et al., 2012; Benedict et al., 2007; Kan et al., 2008; Adams and Kapan, 2009; Saba et al., 2014a, 2014b).

Furthermore, the spatial identification of dengue risk in urban areas, where highest incidence often occurs, is essential to elaborate effective control measures (Wen et al., 2015), and it can be performed more effectively by not overlooking the important role of human mobility (Prothero, 1977; Adams and Kapan, 2009; Stoddard et al., 2009). In particular, it has been suggested that commuters, if they are infected, can play an important role by enhancing the pace of diffusion (Wen et al., 2012), and that the “transport interfaces” (as such a metro station) can be used to better evaluate environmental exposure in epidemiologic research (Perchoux et al., 2013). Despite this increasing interest amidst the aforementioned importance of modern urbanization to the spread of dengue, however, there has been no study on the impact of public metro transportation system on local spread of dengue in urban settings, to our best knowledge. It is most likely due to the difficulty in collecting relevant locality, transportation and disease incidence data with sufficient detail and more importantly, in suitably quantifying the combined effect of proximity and passenger flow.

Straddling the Tropic of Cancer, tropical southern Taiwan has favorable climate conditions for proliferation of *Aedes aegypti*, and consequently dengue outbreaks are quite frequent. The most affected area is Kaohsiung City, which accounted for more than half of all dengue cases reported over the island in the last 18 years (Taiwan CDC, 2016). Moreover, the number of reported dengue cases in Kaohsiung has increased dramatically in the past two years (2014–2015). The July 2014 gas explosion in Kaohsiung had surely contributed greatly to the recording-breaking DENV-1 outbreak in Kaohsiung in 2014 (Wang et al., 2016, 2015). However, there were neither any such catastrophic events nor any unusual climatological event, such as typhoon, that could have led to the even more severe outbreak in Kaohsiung in 2015. Moreover, each of these two massive local outbreaks resulted in more reported cases than the total case number of all previous years in Kaohsiung before 2014. Consequently, question arises as to what factors could have contributed to the local spread of the disease.

Several studies have been conducted to investigate the spatial distribution of dengue cases in Kaohsiung. For examples, to evaluate the influence of population density, proximity to transportation arteries and water bodies (Hsueh et al., 2012), to analyze the geographical heterogeneities in dengue-mosquito and dengue-human relationships (Lin and Wen, 2011), and to identify hot spots (Lin et al., 2012) and risk areas (Wen et al., 2006) for transmission. Moreover, the most densely populated districts near the city central area with multiple Kaohsiung Rapid Transit (KRT) stations, Lingya, Qianzhen, and particularly Sanmin, have also had the highest dengue fever incidence every year in Kaohsiung.

Our aim is to investigate the role of the KRT, with >50 million passengers yearly in recent years (KCGSTAT, 2016), as a possible contributing factor to the spread of dengue in the areas surrounding the KRT system. Utilizing geographic information system (GIS) tool, we examine the geographical distribution of the two epidemics in 2014 and 2015 to evaluate the relationship between dengue incidence rate and proximity to the KRT stations. Moreover, we include in our analysis the passenger traffic information, i.e. the number of people entering and exiting each KRT station. We assume that the impact of the transportation system on the disease spreading process: (i) decreases with distance, but also (ii) varies from one stop to another so that more users imply greater contagion risk.

2. Materials and methods

2.1. Data

In this study, we adopt “Li” (the fourth level administrative subdivision in Taiwan) as the primary unit of observation. We downloaded the polygon shapefile representing the Li's of the special municipality of Kaohsiung, from the geographic database available at the Taiwan Government Open Data Platform website (Taiwan NDC, 2016a). Similarly, we retrieved the point shapefile of the KRT stations from the Ministry of Transportation and Communications (Taiwan MOTC, 2016). The number of dengue fever cases reported in each Li was collected from the Taiwan Government Open Data Platform website (Taiwan NDC, 2016b). We focus on the epidemics that occurred in 2014 and 2015, as they were the most severe epidemics in the last 20 years, totaling 14,970 and 19,704 cases, respectively.

The Civil Affairs Bureau of the Kaohsiung City Government website (KCGCABU, 2016) provides the demographic trend at Li level on a monthly basis. The population data at the end of the years were then retrieved to calculate the incidence rate, defined by the cumulative number of cases per 100,000 inhabitants of each Li from February 15 to the end of the year. The area of each Li was subsequently measured by means of the open source geographic information system QGIS (version 2.14.2-Essen) from official government polygon shapefile of the Li's, in order to obtain the corresponding yearly population density. The monthly traffic volume of each KRT station, i.e. the number of passengers entering and exiting, is available at the Department of Statistics of the Kaohsiung City Government website (KCGSTAT, 2016).

2.2. Proximity and traffic volume: calculation of RAMPP

The KRT currently has two lines: Red line with 24 stations that runs north south and Orange line with 15 stations that runs east west. We focus on the Li's within a 1-km radius from at least one KRT station, approximating the station-Li distance as follows. Using QGIS, we determine the centroid of each Li and create 10 circles (or buffers in GIS terminology) around each station of radius varying from 100 m to 1 km with a 100 meter step. We then identify, for each station, the Li's with centroids contained in each of its 10 annuli (i.e. the ring-shaped regions bounded by two consecutive concentric circles). Finally, we define the station-Li distance d(s,l) as the outer radius of the annulus to which the corresponding centroid belongs. In other words, we consider a radial distance, rounded up to the nearest 100 m.

Since most of the circles overlap, some Li's are located in the proximity of more than one KRT station. Fig. 1 provides an example of the station-Li distance d(s,l). Here two KRT stations belonging to the Red Line are displayed, namely "Kaohsiung Arena" and "Aozihdi", together with six (out of 10) of their circles from 500 m to 1000 m. Five Li's (in yellow) are located within 1 km from at least one of the two stations, namely "Xinzhong", "Huafeng", "Mingcheng", "Xinshang", and "Longzi".

We also include one Li in which a station is located but is 1.01 km from the centroid of the Li ("Hongnan", located in the Nanzi District).

We quantify the proximity and passenger traffic information for each Li, by introducing an index that enables us to quantify the contagion risk associated with the nearby presence of one or more KRT stations. We call this index "Risk Associated with Metro Passengers Presence" or RAMPP for short. RAMPP takes into account, simultaneously, the passenger traffic of all KRT stations located within a fixed radius as well as their actual distance, while giving more weight to the busier and/or closer stations.

We adopt the sum of the monthly traffic volume of each KRT station, recorded from March to December, as the reference datum for each year. Thanks to these data, the RAMPP index is calculated for each Li applying the following formula: $RAMPP(l,y) = \sum_{i=1}^{n_l} \frac{tv(S_i,y)}{d(S_i,l)}$, where $\{S_i\}_l$ is the set of stations falling within 1 km from Li l, n_l is the cardinality of $\{S_i\}_l$, $tv(S_i,y)$ is the traffic volume of S_i over year y, and $d(S_i,l)$ is the distance between S_i and l. It is clear from its definition that the RAMPP: (i) adds up the effect of all the stations within a given radius, (ii) decreases with distance and (iii) increases with traffic volume.

2.3. Statistical analysis

Preliminary, the relationship between yearly incidence rate and RAMPP is evaluated by calculating the Spearman's correlation coefficient, and the result is compared to that obtained between yearly incidence rate and population density.

We cluster the Li's according to their RAMMP value, by applying the K-means method (Hartigan and Wong, 1979), in order to minimize the variability within clusters and maximize the variability between clusters. The within-groups sum of squared errors is then computed to assess the clustering quality for K = 2 to 10. Subsequently, a one-way analysis of variance on ranks (i.e. the Kruskal-Wallis' test) is carried out on the yearly incidence rate data, considering as a factor the grouping provided by the K-means algorithm. Finally, we use the Dunn's test (with Šidák adjustment) as a post hoc test. Statistical analyses are carried out for the two epidemics in 2014 and 2015 using the RStudio software (version 0.99.893).

3. Results

Following the procedure described above and illustrated in Fig. 1, we select 314 Li's (out of a total number of 891 Li's in Kaohsiung). Moreover, for each of these Li's we identify all stations within 1 km of its centroid, again using the previously described distance. The cases reported in these 314 Li's within close proximity of the KRT system, account for almost half of the total case number reported in Kaohsiung in 2014 and 2015, or more precisely, 43.41% (6499 cases) and 46.78% (9219 cases), respectively. For illustration, Fig. 2 provides a geographic map of the 314 Li's with respective incidence rates in 2014 and 2015. The frequency distribution of the Li's by yearly dengue incidence rate in 2014–2015 is also shown in Fig. 3.

We make use of the RAMPP index, in order to differentiate the weight of the KRT stations, based on their traffic volume. Since it is reasonable to suppose that the busiest stations are usually located in densely populated areas, we first compare the correlation between RAMPP and yearly incidence rate with the correlation between population density and yearly incidence rate, for both years. Thus, we

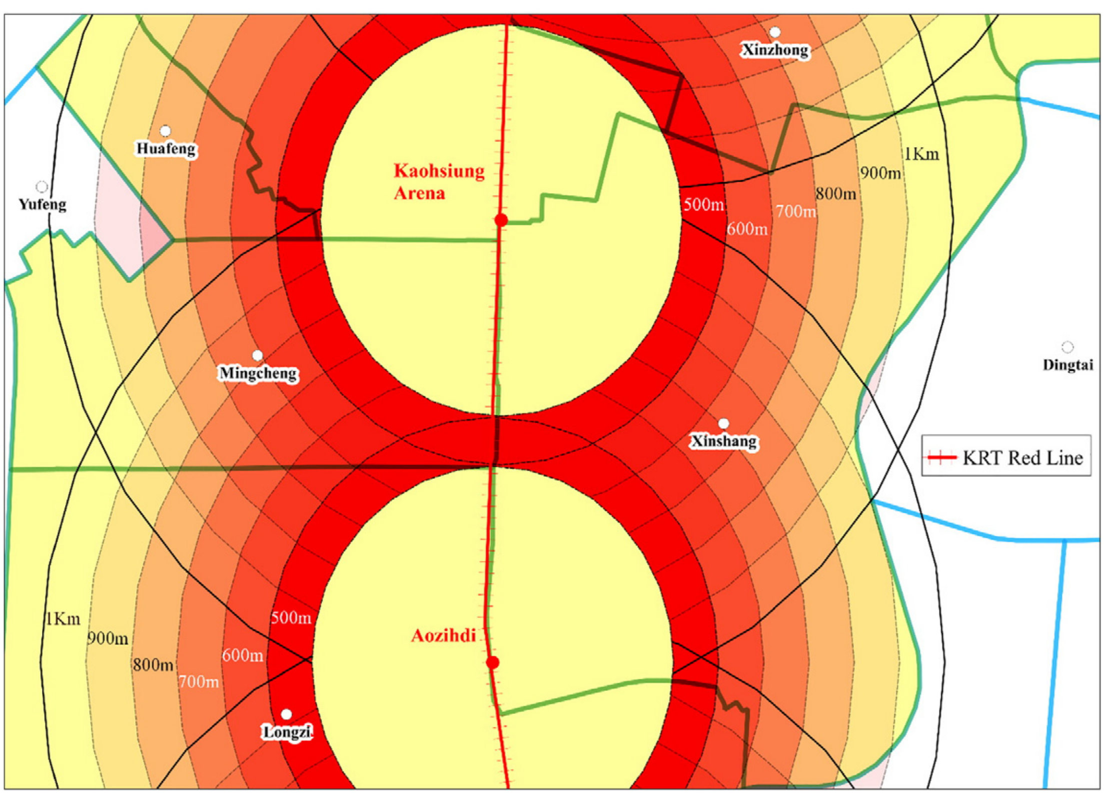


Fig. 1. Example of the station-Li distance d(s,l) of two stations ("Kaohsiung Arena" and "Aozihdi") belonging to the Red Line together with six (out of 10) of their circles from 500 m to 1000 m, in different shades of red. Five Li's ("Xinzhong", "Huafeng", "Mingcheng", "Xinshang" and "Longzi" in yellow) are located within 1 km from at least one of the two stations.

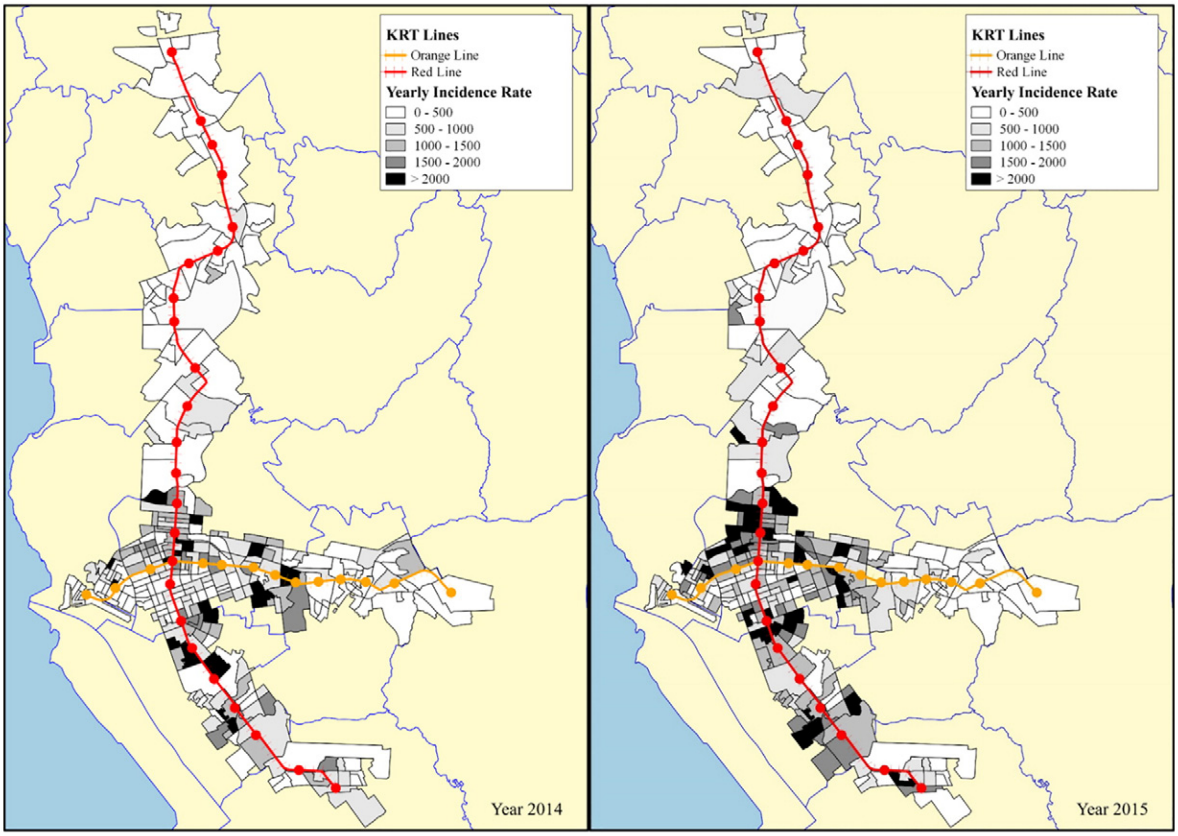


Fig. 2. Geographic map of 314 Li's in Kaohsiung with respective incidence rates in 2014 and 2015.

have four Spearman's coefficients. The result shows that, using RAMPP, the correlation coefficient is higher in both 2014 (0.25 vs. 0.22 – p-values ≪ 0.05) and 2015 (0.26 vs. 0.21 - p-values ≪ 0.05). The K-means method is then applied for K = 2 to 10 on both 2014 and 2015 data, and the within-groups sums of squared errors are displayed in Fig. 4.

Since for K > 4, the effect on the sum of squared errors is negligible, we opt to perform the successive analyses, taking into account three and four clusters, respectively. Table 1 shows the RAMPP mean value for each clustering solution, while the frequency distribution of the Li's with average and standard error of yearly incidence rate for each RAMPP-based cluster is given in Table 2.

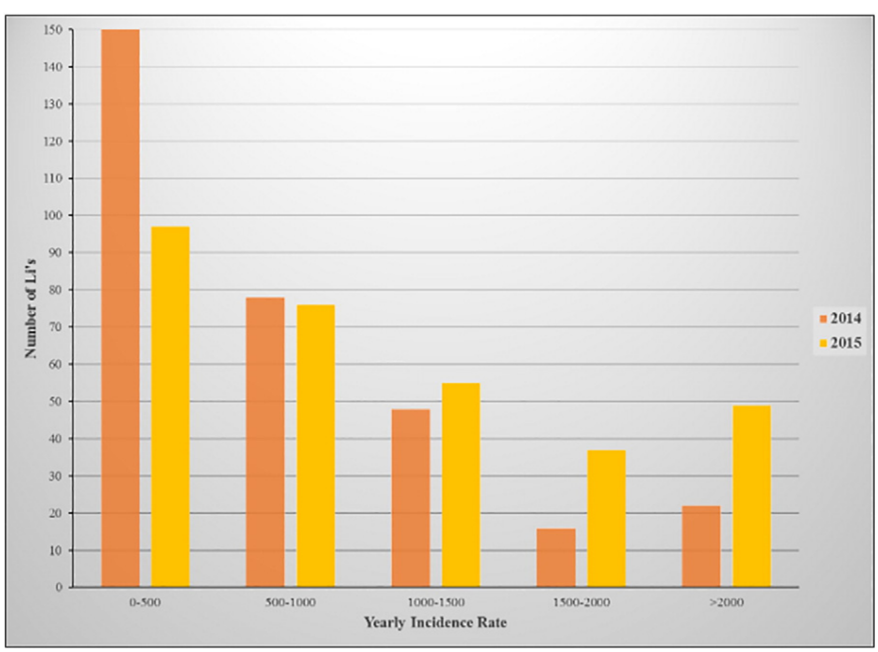


Fig. 3. The frequency distribution of the Li's by yearly dengue incidence rate. Orange bars refer to 2014 data; yellow bars refer to 2015 data.

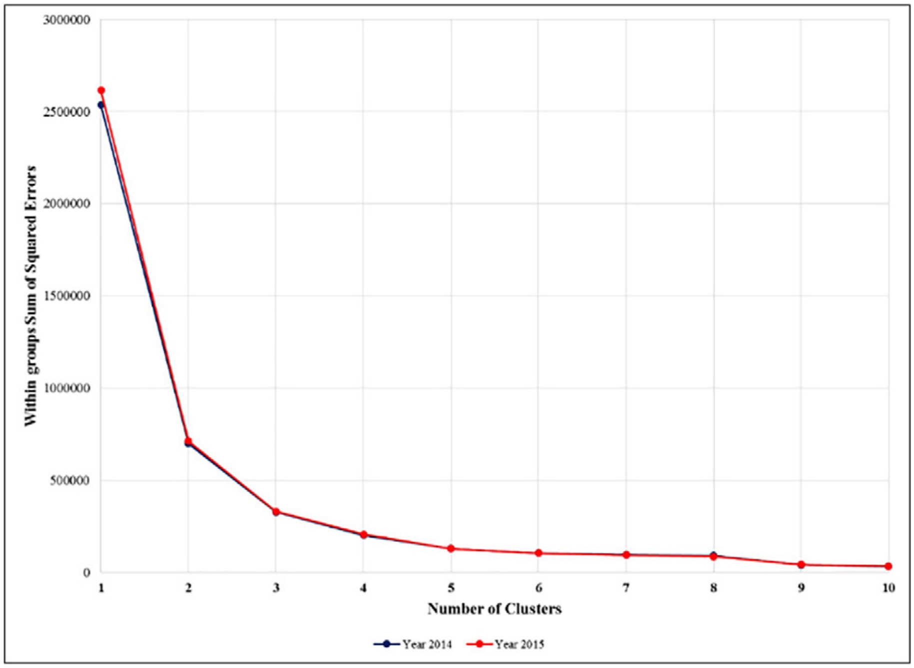


Fig. 4. Evaluation of the K-means clustering: within group sum of squared errors as a function of the number of clusters.

As an illustrative example, for the 2014 epidemic the 314 Li's belonging to four different RAMPP-based clusters are colored in four shades of gray from the darkest shade denoting the cluster with the highest RAMPP mean value, as shown in Fig. 5.

For the case K = 4, two interactive maps are provided as Supplementary Files, displaying simultaneously the yearly incidence rate and the belonging RAMPP-based cluster of each Li.

For both years and for each clustering solution, the normality of data is determined, by applying the Shapiro-Wilk's test on the residuals of the linear model, while the variance homogeneity is assessed via the Levene's test. The data normality requirement is not met (p-values $\ll 0.05$), while the variance homogeneity assumption is largely satisfied (p-values $\gg 0.05$). Subsequently, four Kruskal-Wallis' tests are performed, adopting the RAMPP-based classification as the independent factor, with significant differences detected in both 2014 (p-value $= 4.99\text{E}-05$ for K = 3, p-value $= 3.47\text{E}-05$ for K = 4) and 2015 datasets (p-value $= 5.51\text{E}-05$ for K = 3, p-value $= 2.90\text{E}-04$ for K = 4).

Consequently, it is possible to apply the Dunn's post hoc test (with Šidák adjustment), in order to compare the group means, as given in Fig. 6 for 2014–2015.

Considering K = 3, the outcome of the analysis conducted on the 2014 data shows that the (average) yearly incidence rate observed in Cluster 3, i.e. the cluster collecting the Li's with the highest RAMPP values, is always significantly greater than that recorded in the clusters with small RAMPP values (Table 3). In 2015, no significant difference is detected between Cluster 3 and Cluster 2, but both are significantly higher than Cluster 1 (Table 4).

For K = 4, the Dunn's test shows again that Cluster 4, i.e. the cluster collecting the Li's with the highest RAMPP values, in 2014 exhibits an yearly incidence rate significantly greater than that recorded in both Cluster 1 and Cluster 2. A significant difference is also detected between Cluster 3 and Cluster 1. Moreover, no significant difference can be detected between Cluster 4 and Cluster 3, between Cluster 3 and Cluster 2, and between Cluster 2 and Cluster 1 (Table 5). Similarly, for the results obtained for the 2015 data (Table 6).

4. Discussion

While the relationship between dengue incidence and population density is well established (Murray et al., 2013), our use of RAMPP takes into account the population mobility in an urban setting via the use of passenger traffic information of a public transportation system, providing a new perspective in identifying high risk areas for dengue

Table 1
RAMPP mean values of the clusters identified by the K-means algorithm for K = 3 and K = 4.

K	Cluster number	RAMPP mean value	
		Year 2014	Year 2015
3	1	4736.37	4530.20
	2	15,786.94	15,603.40
	3	30,483.59	30,058.79
4	1	3285.49	3165.45
	2	8778.27	8577.08
	3	17,931.20	17,803.79
	4	31,667.23	31,235.74

Table 2
The frequency distribution of the Li's with average and standard error of yearly incidence rate, for each RAMPP-based cluster.

Year	K	Cluster number	# of Li's	Average yearly incidence rate	Standard error
2014	3	1	209	688.80	48.93
		2	71	779.14	81.44
		3	34	1115.24	109.58
	4	1	143	639.18	55.80
		2	86	749.59	79.56
		3	56	855.25	92.59
		4	29	1152.95	126.07
2015	3	1	208	1000.33	65.67
		2	72	1333.93	128.77
		3	34	1433.41	113.89
	4	1	144	982.54	78.95
		2	85	1083.57	101.00
		3	56	1389.90	153.40
		4	29	1428.41	117.88

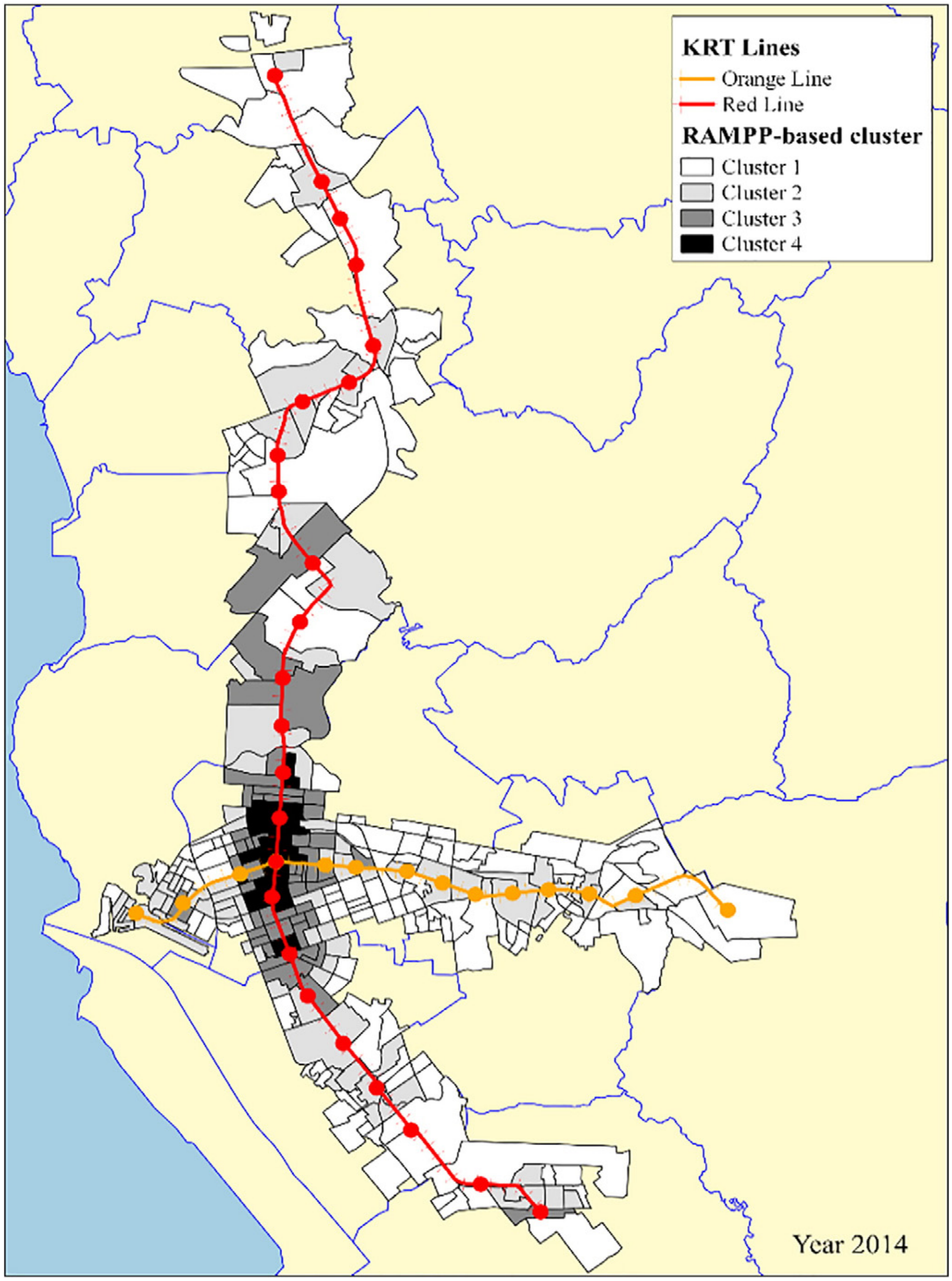


Fig. 5. Geographic map of 314 Li's in Kaohsiung. Based on 2014 data, the 314 Li's in four different RAMPP-based clusters are colored in four shades of gray from the darkest shade denoting the cluster with the highest RAMPP mean value.

transmission. The RAMPP index reflects the population density, as the location of the stations is planned to provide easy access to as many people locally as possible, and consequently the most crowded stations are often located in highly populated areas. At the same time, however, our index carries additional and more "dynamic" information. *Aedes* mosquitos (*Ae. Aegypti* and *Ae. albopictus*) are known to be diurnal feeders (Carrington et al., 2013; Farajollahi et al., 2012). It is exactly during the daytime that people living in proximity of one or more busy stations can come into contact with a number of individuals (students, workers, etc.), which is potentially even larger than the population itself, thus increasing the infection probability. In other words, the RAMPP index somehow incorporates and extends the information related to the population density, and such consideration is consistent with the fact that the correlation between RAMPP and dengue incidence is greater than the correlation between population density and dengue incidence.

By examining the geographic pattern of yearly incidence rate in 2014 (Fig. 2), it can be observed that all Li's with at least 1000 infections per 100,000 inhabitants are located in the central area in the districts of Yancheng, Sanmin, Xinxing, Lingya, Qianzhen, and Fengshan. The only exceptions are two Li's in Xiaogang district, close to Kaohsiung International Airport, and one Li in Nanzi district. It is interesting to note that most of these Li's are located along the southern branch of the Red Line (from Houyi station to Siaogang station) or the central-western branch of the Orange Line (from Yanchengpu station to Weiwuying station). One can make similar observations regarding the 2015 epidemic (Fig. 2). Due to the larger amount of cases recorded, the number of Li's with an elevated incidence rate obviously increases in 2015; however,

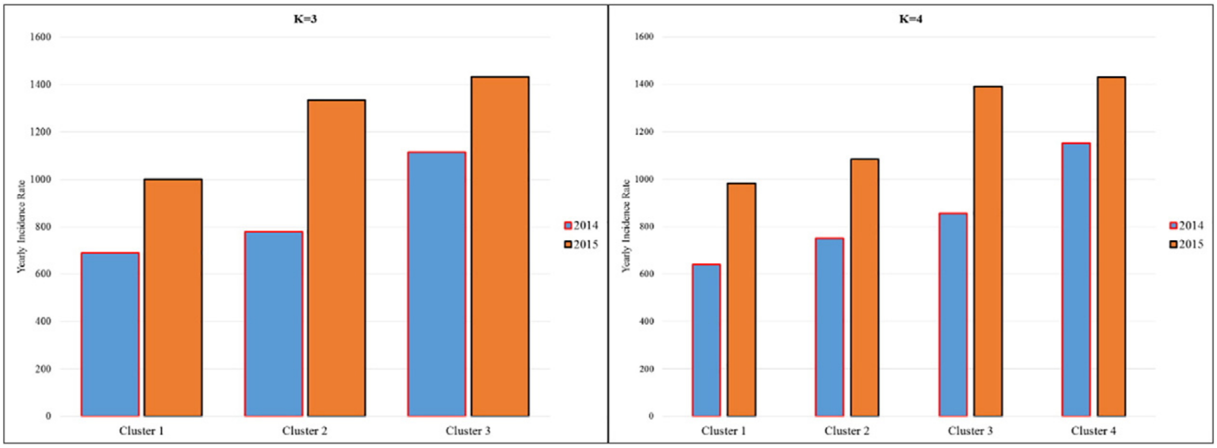


Fig. 6. Mean value of yearly dengue incidence rate by RAMPP-based cluster, for K = 3 (on the left) and for K = 4 (on the right). Blue bars with a red border refer to 2014 data; orange bars with a black border refer to 2015 data.

their spatial distribution at district level is almost the same. With respect to 2014, only a couple of Li's with incidence rate > 1500 are registered in both Gushan and Zuoying district.

The geographic patterns of the RAMPP index in 2014 and in 2015 (Fig. 5) are clearly almost identical, as the number of passengers does not change much from one year to another. Considering the four cluster solution, we note that the Li's belonging to Cluster 4 are located around the following six stations: Houyi, Kaohsiung Main Station, Formosa Boulevard, City Council, Central Park, and Sanduo Shopping District. Moreover, districts that include at least one Li belonging to Cluster 3 or Cluster 4 are Zuoying, Gushan, Sanmin, Yancheng, Qianjin, Xinxing, Lingya, Qianzhen, and Xiaogang. Except for Qianjin, all these districts have been cited in the previous paragraph as the districts including Li's with the highest level of yearly incidence rate.

Naturally, the areas with the highest incidence do not perfectly correspond to the ones with the highest RAMPP, because a perfect match would require the existence of a relationship between dengue incidence and RAMPP much tighter than the one emerged from this study. The Dunn's tests performed on the RAMPP-based clusters show that, on average, the yearly incidence rate is significantly higher, where the risk associated with the presence of KRT stations is elevated. In fact, despite the number of cases reported during the 2014 outbreak was lower than that in 2015, and considering two different clustering solutions, the statistical analyses produce largely consistent results (Tables 3–6), showing that the Li's with the highest RAMMP exhibit (on average) the highest yearly incidence rate. Moreover, for both epidemics, the behavior of the average incidence rate is monotonic, with respect to the RAMPP-based clustering (Fig. 6), corroborating the idea that the RAMMP index that we propose can be effectively employed as criterion for the identification of high/low risk areas.

The evaluation of the impact of the KRT system on the spreading of dengue in Kaohsiung highlights substantial disparity in the role KRT stations play in disease diffusion, and that actual passenger volume of each station is a key to obtaining reasonable results. During these two years, the KRT subway lines served around 60 million passengers per year, but the distribution of traffic volume was far from homogeneous. Indeed, the two busiest KRT stations in Kaohsiung, namely “Kaohsiung Main Station” and “Zuoying” (the latter of which links the KRT with the Taiwan High Speed Rail System), accounted for >10 million passengers each per year. While the two least frequented ones (“Qingpu” and “Qiaotou Sugar Refinery”) had around 650,000 passengers (data source: KCGSTAT, 2016). The RAMPP serves the purpose of capturing such “influence variation”, reflecting the simple but important idea that a large amount of passenger flow in and out of a station can dramatically increase the contagion risk in the neighborhoods. While on the other hand, a sporadically used station is unlikely to be able to substantially impact the spreading process.

Our findings could have important public health implications. In the aftermath of 2003 Severe Acute Respiratory Syndrome (SARS) epidemic, several countries such as Singapore and Taiwan introduced non-contact infrared thermometers (NCIT) at border control such as airports in order to detect febrile passengers, in an attempt to detect imported dengue cases or to delay the introduction of a novel influenza strain (Bitar et al., 2009). A recent study in Taiwan (Chang et al., 2016) indicates that in recent years, close to half of the imported dengue cases in Taiwan were detected at border every year (Fig. 7), providing evidence for its effectiveness in detection of early cases. Epidemiologic investigation by Taiwan CDC indicates that almost 98% of the reported cases in Kaohsiung were infected locally in their Li of residence. Our study, showing the significant role that mass transportation system could play in dissemination of disease, would suggest that establishing similar NCIT sites at the busiest KRT stations could be very effective in early detection of dengue cases, and subsequently reduce local transmissions in the nearby communities.

This study certainly has its limitations, which need to be noted. First, the dataset used to evaluate the proposed methodology pertains only to one city, Kaohsiung, mainly because it is difficult to gather locality, mass transportation, and disease incidence data with the required detail level. Moreover, analysis was performed for two epidemics (in 2014 and 2015), because the KRT system only began to operate in 2008, and

Table 3
Year 2014 - Dunn's test result for K = 3: pairwise comparison of yearly incidence rate by RAMPP-based cluster (with Šidák adjustment).

Column mean - row mean *p*-Value	Cluster 3	Cluster 1
Cluster 1	426.44 0.0000*	
Cluster 2	336.09 0.0092*	−90.34 0.1253

* p-value < 0.05.

Table 4
Year 2015 - Dunn's test result for K = 3: pairwise comparison of yearly incidence rate by RAMPP-based cluster (with Šidák adjustment).

Column mean - row mean *p*-Value	Cluster 3	Cluster 1
Cluster 1	433.08 0.0002*	
Cluster 2	99.48 0.1948	−333.60 0.0052*

* p-value < 0.05.

Table 5
Year 2014 - Dunn's test result for K = 4: pairwise comparison of yearly incidence rate by RAMPP-based cluster (with Šidák adjustment).

Column mean - row mean *p*-Value	Cluster 3	Cluster 1	Cluster 2
Cluster 1	216.07 0.0101*		
Cluster 2	105.66 0.2961	−110.41 0.3988	
Cluster 4	−297.70 0.1700	−513.77 0.0000*	−403.36 0.0033*

* *p*-value < 0.05.

during 2008–2013 no major dengue outbreaks occurred in Kaohsiung. Secondly, the K-means method adopted to cluster the Li's based on their RAMPP value is strictly statistical. Further applications of the RAMPP index could conceivably help to better define its properties, and consequently resulting in a less “aseptic” clustering procedure.

Moreover, a number of confounding factors could have influenced the results outlined here. In addition to the role of population density, which has already been addressed, the presence of popular meeting places (shopping malls, tourist spots, schools, etc.) or potential mosquito breeding sites in the vicinity of the KRT stations could have increased the infection rate in that areas, more than the traffic volume. However, unlike traffic volume, it is very difficult to quantify the number of people actually visiting those places (even on a yearly basis), and consequently the transmission risk. Furthermore, it is sensible to assume that the KRT stations are often located in the proximity of popular meeting places (or vice versa). Therefore, the RAMPP index, taking into account of the distance, could implicitly include this information as well as population density. In any case, the approach adopted in this paper is strictly empirical, as our aim is to highlight the statistically based relationship detected between dengue incidence and the RAMPP index, which can certainly be exploited for practical purposes by public health decision-makers.

5. Conclusion

The diffusion of an infectious disease is a very complex phenomenon, depending on a huge number of factors. Public transportation system is just one of them and its weight can vary from one outbreak to another, as every year the conditions could change, magnifying or diminishing the impact of all the variables at play. Consequently, it would be unrealistic to expect from a straightforward quantifier, such as the proposed RAMPP, to be able to capture fully the intricate variability of each epidemic under any conditions.

Despite these aforementioned limitations, the close similarity of the results obtained from data of two different years is a promising indication of the usefulness of our proposed index, and provides motivation to investigate its applicability through further studies. Enhancing our understanding of how public rapid transit system contributes to disease spread in densely populated urban areas, could be useful in the design of more effective and timely intervention and control measures for future outbreaks.

Table 6
Year 2015 - Dunn's test result for K = 4: pairwise comparison of yearly incidence rate by RAMPP-based cluster (with Šidák adjustment).

Column mean - row mean *p*-Value	Cluster 3	Cluster 1	Cluster 2
Cluster 1	407.36 0.0078*		
Cluster 2	306.33 0.2326	−101.03 0.4514	
Cluster 4	−38.51 0.5081	−445.87 0.0006*	−344.84 0.0226*

* *p*-value < 0.05.

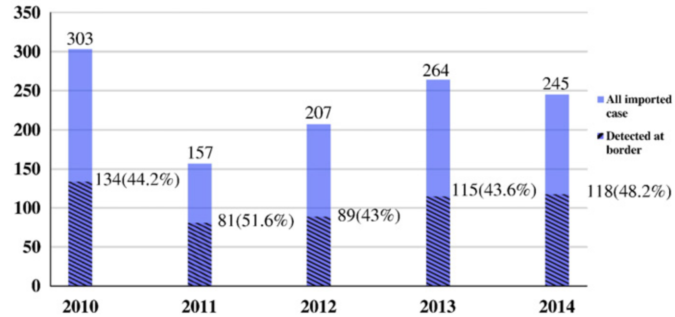


Fig. 7. Yearly number of imported dengue cases and the number (percentage) detected at border in Taiwan, 2010–2014.
(Source: Chang et al., 2016).

Conflict of interest

None.

Acknowledgment

MS is supported by a Taiwan Ministry of Science and Technology postdoctoral fellowship (104-2811-B-039-005). YHH is supported by funding from Taiwan Ministry of Science and Technology grants (103-2314-B-039-010-MY3, 103-2115-M-039-002-MY2). The authors are grateful to the reviewer and the editor for their constructive comments, which greatly improve this manuscript.

Appendix A. Supplementary data

The files 2014.kmz and 2015.kmz display simultaneously the yearly incidence rate and the RAMPP-based cluster each Li belongs to, for the case K = 4. They represent, for each year, the yearly incidence rate class with different colors, and the RAMPP-based cluster by “lifting” each element at four different altitudes. Please note that for a correct 3D visualization, the files must be downloaded in a local folder and opened with Google Earth, which needs to be installed. Supplementary data associated with this article can be found in the online version, at http://dx.doi.org/10.1016/j.scitotenv.2017.04.050. These data include the Google Maps of the most important areas described in this article.

References

Adams, B., Kapan, D.D., 2009. Man bites mosquito: understanding the contribution of human movement to vector-borne disease dynamics. PLoS One 4, e6763. http://dx.doi.org/10.1371/journal.pone.0006763.
Barrett, C.L., Bisset, K.R., Eubank, S.G., Feng, X., Marathe, M.V., 2008. EpiSimdemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks. Proceedings of the 2008 ACM/IEEE Conference on Supercomputing 2008. IEEE Press, p. 37.
Benedict, M.Q., Levine, R.S., Hawley, W.A., Lounibos, L.P., 2007. Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. Vector Borne Zoonotic Dis. 7:76–85. http://dx.doi.org/10.1089/vbz.2006.0562.
Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., Myers, M.F., 2013. The global distribution and burden of dengue. Nature 496 (7446):504–507. http://dx.doi.org/10.1038/nature12060.
Bitar, D., Goubar, A., Desenclos, J.C., 2009. International travels and fever screening during epidemics: a literature review on the effectiveness and potential use of non-contact infrared thermometers. Euro Surveill. 14 (6), 19115.
Carrington, L.B., Seifert, S.N., Willits, N.H., Lambrechts, L., Scott, T.W., 2013. Large diurnal temperature fluctuations negatively influence *Aedes aegypti* (Diptera: Culicidae) life-history traits. J. Med. Entomol. 50 (1), 43–51.
Chang, Y.P., Jhao, C.C., Wu, Y.C., Ho, L.L., Su, C.F., 2016. Review of simplifying border quarantine practice in Taiwan. Taiwan Epidemiol. Bull. 32 (24):514–524. http://dx.doi.org/10.6525/TEB.20161220.32(24).001 (In Chinese).

Colizza, V., Barrat, A., Barthélemy, M., Vespignan, A., 2007. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. BMC Med. 5:34. http://dx.doi.org/10.1186/1741-7015-5-34.

Edelson, P.J., Phypers, M., 2011. TB transmission on public transportation: a review of published studies and recommendations for contact tracing. Travel Med. Infect. Dis. 9:27–31. http://dx.doi.org/10.1016/j.tmaid.2010.11.001.

Farajollahi, A., Healy, S.P., Unlu, I., Gaugler, R., Fonseca, D.M., 2012. Effectiveness of ultra-low volume nighttime applications of an adulticide against diurnal *Aedes albopictus*, a critical vector of dengue and chikungunya viruses. PLoS One 7 (11), e49181.

Gubler, D.J., 2011. Dengue, urbanization and globalization: the unholy trinity of the 21st century. Trop. Med. Health. 39:S3–S11. http://dx.doi.org/10.2149/tmh.2011-S05.

Hartigan, J.A., Wong, M.A., 1979. Algorithm AS 136: a k-means clustering algorithm. J. R. Stat. Soc.: Ser. C: Appl. Stat. 28 (1):100–108. http://dx.doi.org/10.2307/2346830.

Horna-Campos, O.J., Sánchez-Pérez, H.J., Sánchez, I., Bedoya, A., Martín, M., 2007. Public transportation and pulmonary tuberculosis, Lima, Peru. Emerg. Infect. Dis. 13, 1491–1493.

Hsieh, Y.H., van den Driessche, P., Wang, L., 2007. Impact of travel between patches for spatial spread of disease. Bull. Math. Biol. 69:1355–1375. http://dx.doi.org/10.1007/s11538-006-9169-6.

Hsieh, Y.H., de Arazoza, H., Lounes, R., 2013. Temporal trends and regional variability of 2001–2002 multiwave DENV-3 epidemic in Havana City: did Hurricane Michelle contribute to its severity? Tropical Med. Int. Health 18:830–838. http://dx.doi.org/10.1111/tmi.12105.

Hsueh, Y.H., Lee, J., Beltz, L., 2012. Spatio-temporal patterns of dengue fever cases in Kaohsiung City, Taiwan, 2003–2008. Appl. Geogr. 2012 (34):587–594. http://dx.doi.org/10.1016/j.apgeog.2012.03.003.

Kan, C.C., Lee, P.F., Wen, T.H., Chao, D.Y., Wu, M.H., Lin, N.H., Huang, S.Y.J., Shang, C.S., Fan, I.C., Shu, P.Y., Huang, J.H., 2008. Two clustering diffusion patterns identified from the 2001–2003 dengue epidemic, Kaohsiung, Taiwan. Am.J.Trop. Med. Hyg. 79 (3), 344–352.

KCGCABU - Civil Affairs Bureau, Kaohsiung City Government, 2016. Statistics and data: demographics. Demographic statistics survey: monthly number of households in by districts and Li's in Kaohsiung. http://cabu.kcg.gov.tw/Web/StatRpts/StatRpt1.aspx (accessed June 2016). (In Chinese).

KCGSTAT - Department of Statistics, Kaohsiung City Government, 2016. Official statistics and information database. http://kcgdg.kcg.gov.tw/kcgstat/page/kcg08.aspx (accessed June 2016). (In Chinese).

Khan, K., Arino, J., Hu, W., Raposo, P., Sears, J., Calderon, F., Heidebrecht, C., Macdonald, M., Liauw, J., Chan, A., Gardam, M., 2009. Spread of a novel influenza A (H1N1) virus via global airline transportation. N. Engl. J. Med. 361 (2), 212–214.

Lin, C.H., Wen, T.H., 2011. Using geographically weighted regression (GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence of dengue. Int. J. Environ. Res. Public Health 8:2798–2815. http://dx.doi.org/10.3390/ijerph8072798.

Lin, C.H., Schiøler, K.L., Jepsen, M.R., Ho, C.K., Li, S.H., Konradsen, F., 2012. Dengue outbreaks in high-income area, Kaohsiung City, Taiwan, 2003–2009. Emerg. Infect. Dis. 18: 1603–1611. http://dx.doi.org/10.3201/eid1810.111929.

Murray, N.E., Quam, M.B., Wilder-Smith, A., 2013. Epidemiology of dengue: past, present and future prospects. Clin. Epidemiol. 5:299–309. http://dx.doi.org/10.2147/CLEP.S34440.

Perchoux, C., Chaix, B., Cummins, S., Kestens, Y., 2013. Conceptualization and measurement of environmental exposure in epidemiology: accounting for activity space related to daily mobility. Health Place 21:86–93. http://dx.doi.org/10.1016/j.healthplace.2013.01.005.

Prothero, R.M., 1977. Disease and mobility: a neglected factor in epidemiology. Int. J. Epidemiol. 6:259–267. http://dx.doi.org/10.1093/ije/6.3.259.

Saba, H., Miranda, J.G.V., de Freitas Jorge, E.M., Moret, M.A., 2014a. Correlation between transport and occurrence of dengue cases in Bahia. J. Commun. Comput. 11, 217–221.

Saba, H., Miranda, J.G.V., Moret, M.A., 2014b. Self-organized critical phenomenon as a q-exponential decay—avalanche epidemiology of dengue. Phys. A Stat. Mech. Appl. 413:205–211. http://dx.doi.org/10.1016/j.physa.2014.06.045.

Stoddard, S.T., Morrison, A.C., Vazquez-Prokopec, G.M., Soldan, V.P., Kochel, T.J., Kitron, U., Elder, J.P., Scott, T.W., 2009. The role of human movement in the transmission of vector-borne pathogens. PLoS Negl. Trop. Dis. 3 (7):e481. http://dx.doi.org/10.1371/journal.pntd.0000481.

Taiwan CDC - Taiwan Centers for Disease Control, 2016. Taiwan National Infectious Disease Statistics System. http://nidss.cdc.gov.tw/en/ (accessed June 2016).

Taiwan MOTC - Taiwan Ministry of Transportation and Communications, 2016. GIS-T transport network geographic information storage system. Map and data download service. https://gist.motc.gov.tw/MapDataService/Retrieval (accessed May 2016). (In Chinese).

Taiwan NDC – Taiwan National Development Council, 2016a. Government open data platform. Geographic information dataset list. http://data.gov.tw/geodata_shp (accessed January 2016). (In Chinese).

Taiwan NDC – Taiwan National Development Council, 2016b. Government open data platform. Daily confirmed dengue case statistic, 1998. http://data.gov.tw/node/21025 (accessed April 2016). (In Chinese).

Tatem, A.J., Rogers, D.J., Hay, S.I., 2006. Global transport networks and infectious disease spread. Adv. Parasitol. 62:293–343. http://dx.doi.org/10.1016/S0065-308X(05)62009-X.

Wang, S.F., Chang, K., Lu, R.W., Wang, W.H., Chen, Y.H., Chen, M., Wu, D.C., Chen, Y.M.A., 2015. Large dengue virus type 1 outbreak in Taiwan. Emerg. Microbes Infect. 4 (8): e46. http://dx.doi.org/10.1038/emi.2015.46.

Wang, S.F., Wang, W.H., Chang, K., Chen, Y.H., Tseng, S.P., Yen, C.H., Wu, D.C., Chen, Y.M.A., 2016. Severe dengue fever outbreak in Taiwan. Am.J.Trop. Med. Hyg. 94 (1):193–197. http://dx.doi.org/10.4269/ajtmh.15-0422.

Wen, T.H., Lin, N.H., Lin, C.H., King, C.C., Su, M.D., 2006. Spatial mapping of temporal risk characteristics to improve environmental health risk identification: a case study of a dengue epidemic in Taiwan. Sci. Total Environ. 367:631–640. http://dx.doi.org/10.1016/j.scitotenv.2006.02.009.

Wen, T.H., Lin, M.H., Fang, C.T., 2012. Population movement and vector-borne disease transmission: differentiating spatial–temporal diffusion patterns of commuting and noncommuting dengue cases. Ann. Assoc. Am. Geogr. 102:1026–1037. http://dx.doi.org/10.1080/00045608.2012.671130.

Wen, T.H., Lin, M.H., Teng, H.J., Chang, N.T., 2015. Incorporating the human-*Aedes* mosquito interactions into measuring the spatial risk of urban dengue fever. Appl. Geogr. 62: 256–266. http://dx.doi.org/10.1016/j.apgeog.2015.05.003.

WHO, TDR - World Health Organization, Special Programme for Research and Training in Tropical Diseases, 2009. Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control. World Health Organization, Geneva.

Wilson, M.E., 1995. Travel and the emergence of infectious diseases. Emerg. Infect. Dis. 1, 39–46.

Xu, F., McCluskey, C.C., Cressman, R., 2013. Spatial spread of an epidemic through public transportation systems with a hub. Math. Biosci. 246:164–175. http://dx.doi.org/10.1016/j.mbs.2013.08.014.