

# Applying space-time information to explore disease processes

The dynamic patterns of Dengue Fever in Kaohsiung City, 1998-2015.

Wei-Chien-Benny Chin\*, Tzai-Hung Wen, Clive E. Sabel

Lab for Geospatial Computational Science, Department of Geography, National Taiwan University.

\*wcchin.88@gmail.com

Poster-ID: A033

Conference: SRATaiwan 2017, Taichung City



## Problem

Dengue Fever (DF) data is recorded as a space-time point data, i.e. each record contains a spatial location, and a temporal location. Previous studies often use clustering analysis method to analyze this type of disease case data, to search for spatial or space-time cluster areas (hot-spots), and connect the clusters manually. The process between disease clusters has been neglected in previous studies. In **disease diffusion processes**, transmissions may experience a temporal lag for an incubation period, that is the time between infection and disease emergence. Thus a temporal lag between the transmission pairs should be considered in the understanding of disease diffusion [1]. This can help to systematically model the disease diffusion process, and construct the relationships between cases and detect the process chains between clusters (Figure 1). Therefore, the aim of this case study is to detect sub-clusters from the annual DF cases and identify processes between the sub-clusters.

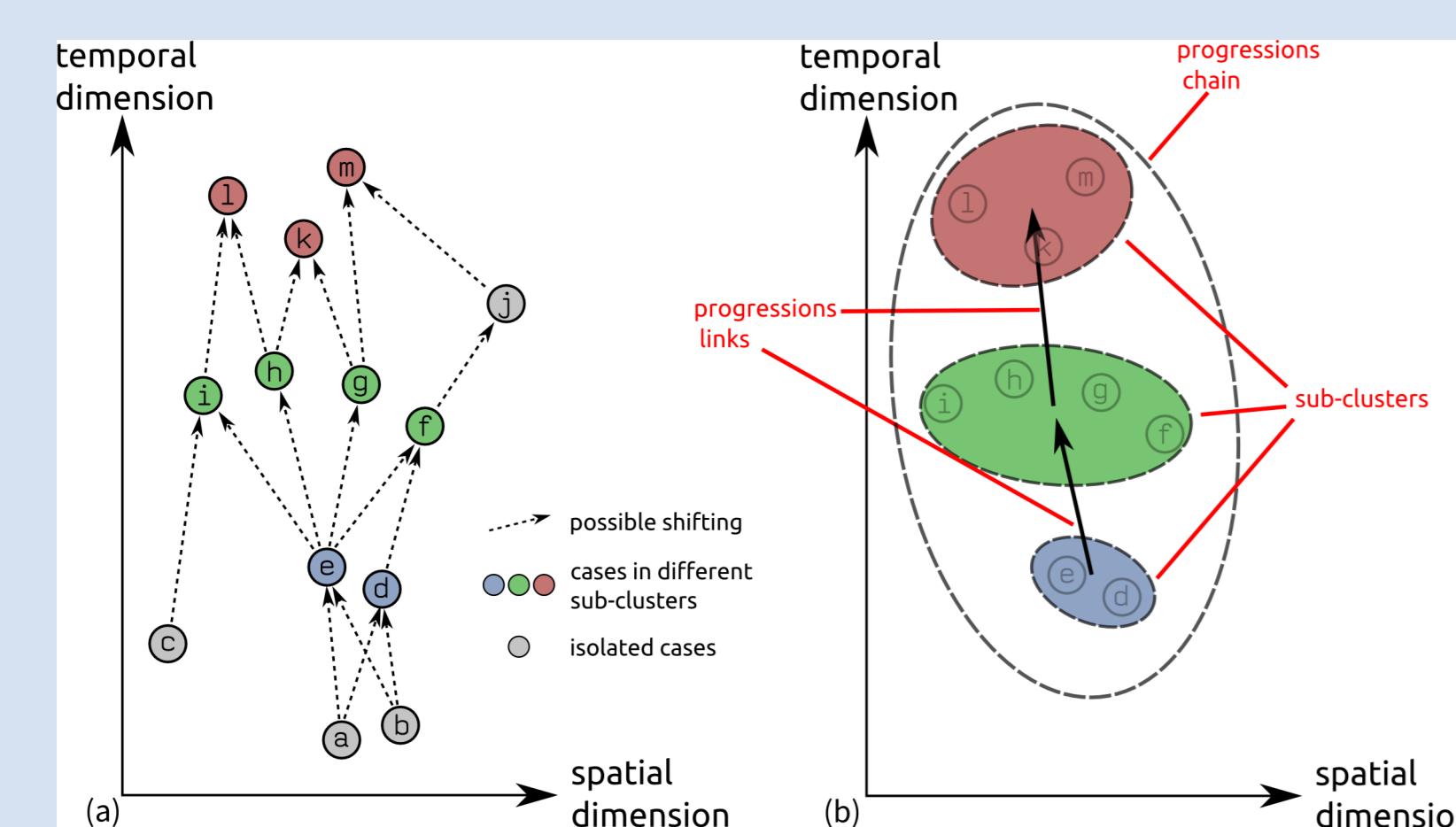


Figure 1: From the cases to sub-clusters, progression links, and progression chains.

## Materials

Our data is provided by Taiwan Centers for Disease Control (Taiwan CDC), which records the daily number of DF cases from 1998 to 2015 in each basic statistical unit (BSU) in Kaohsiung City, separated into imported (to Taiwan) and local indigenous cases, based on the epidemiological investigation records. Only the data of indigenous cases was used in this study to eliminate the external importation noise from the local diffusion process. Each BSU has on average 400 people for the whole of Taiwan. The data is available from Taiwan CDC, under *Taiwan Open Government Data License, version 1.0*. The dataset is available in the Daily reported dengue fever cases since 1998 repository (<http://data.gov.tw/node/21025>).

DF is a vector-borne disease, where the transmission process has a human-mosquito-human cycle. There is a temporal lag between the time a case is infected and the time when the case becomes infectious. The infectious period after the first symptoms appear is 4-5 days; the extrinsic incubation period (EIP) for mosquitoes is 8-12 days; and the intrinsic incubation period (IIP) for humans is 4-10 days [2]. Therefore, the minimum temporal lag between a pair of related cases is 12 days, whereas the maximum time-lag is about 27 days (Figure 2).

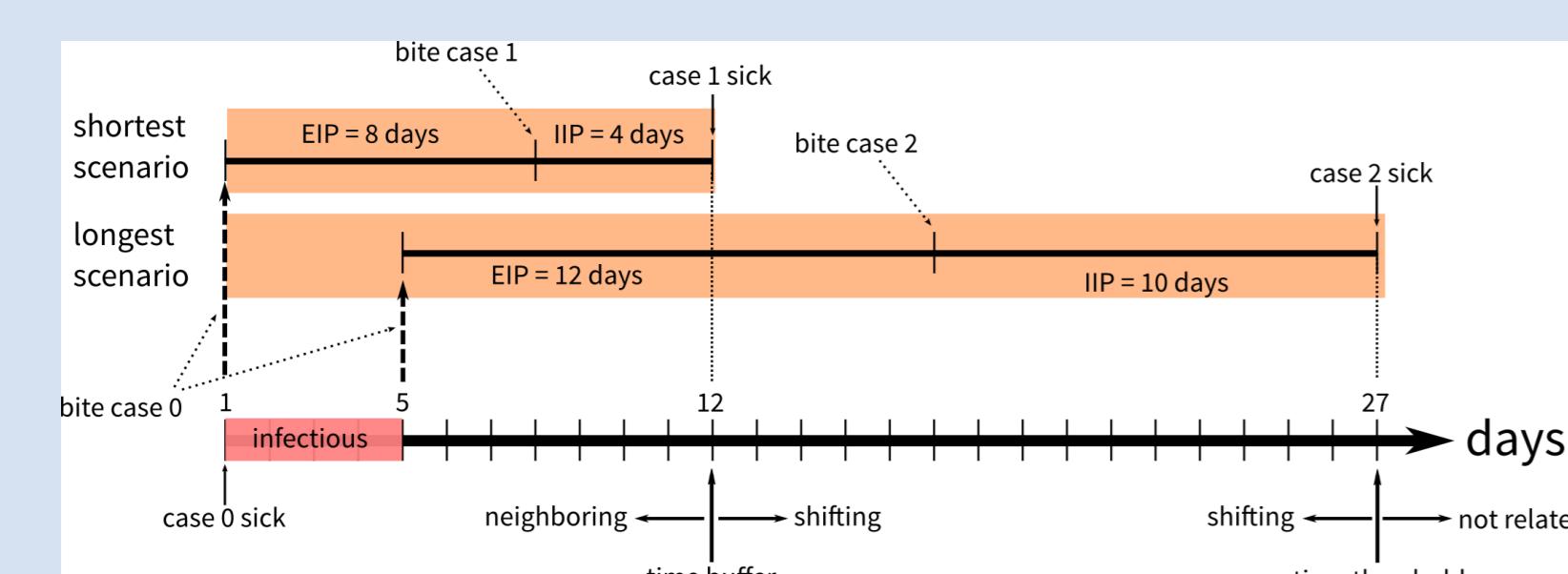


Figure 2: The time parameters settings were designed to capture the ranges of infectious period, EIP, and IIP of Dengue Fever.

## The TaPiTaS algorithm

Our *Tracking Progression In Time And Space* (TaPiTaS) algorithm uses the spatial and temporal distance between each pair of points to identify the most probable common origin and to detect sub-clusters. A sub-cluster is formed by a group of spatially and temporally close points that are probably related to one or several common origins. By common origin, we mean the source or in an epidemiological context, the original infective agent or individual that is common to all subsequent sub-clusters. Then, the progression links between sub-clusters and the progression chains that are formed by several linked sub-clusters could be revealed in the data exploration of the point diffusion process. The illustration of the sub-clusters is shown in Figure 1.

TaPiTaS algorithm was composed of three steps. The first step distinguished the relationships of each pair of the spatially close events into two types: shifting link or neighboring pair. The second step focused on identifying space-time sub-clusters. The third step aimed to construct the progressions between sub-clusters.

## Descriptive statistics of the progressions

The SC size represents the median number of cases within a sub-cluster; chain size represents the median number of sub-clusters within a chain; the SC duration represents the median temporal duration of the sub-clusters, which measures the number of days between the earliest and the latest cases within a sub-cluster; the chain duration represents the median temporal duration of the chains. The numbers in brackets are the median absolute deviation (MAD) of the corresponding columns. MAD is the abbreviation for median absolute deviation, which better describes our range than the normal standard deviation.

The total number of cases varied between years: six out of the 18 years have less than 100 cases; in 2014 and 2015, the number of cases exceeded 10,000. The number of sub-clusters are related to the number of cases in the year, but are not always proportional to the number of cases. For example, 2010 had less cases than 2011, but more sub-clusters were detected. Regardless of the significant differences in the total number of cases, the median number and the MAD of the sub-cluster sizes and durations over the 18 years are similar: sub-clusters consist of 2 to 4 cases, and the duration median is 5 days (with 4 days MAD).

The chain size is measured by using the number of sub-clusters in each chain to show the extent of chains in the year; the chain duration is measured by using the temporal difference between the first case and the last case in each chain to show the temporal continuity in each year. Similar to the sub-cluster sizes and durations, the chain sizes and durations are also similar through the 18 years regardless of the differences in terms of the total size of cases. The chain sizes consist of 2 to 3 sub-clusters and the duration median is one month (with 2 weeks MAD).

Table 1: The descriptive statistics of the diffusion progression from 1998 to 2015, including cases, sub-clusters(SC), progression links(PL), and chains.

year	no. cases	no. SC	SC size	SC duration	no. iso-SC	no. PL	no. chains	chain size	chain duration
1998	113	5	3 (0)	5 (4)	2	2	1	3.0 (0)	66.0 (0)
1999	20	1	-	-	1	0	0	-	-
2000	-	-	-	-	-	-	-	-	-
2001	220	9	8 (9)	14 (15)	3	8	1	6 (0)	81 (0)
2002	4671	276	3 (1)	7 (6)	52	198	39	2 (0)	36 (22)
2003	34	2	2 (0)	2 (2)	2	0	0	-	-
2004	56	1	-	-	1	0	0	-	-
2005	96	6	4 (1)	6 (4)	4	1	1	2 (0)	19 (0)
2006	955	102	3 (1)	4 (4)	21	57	25	3 (1)	29 (18)
2007	168	7	4 (1)	12 (9)	3	2	2	2 (0)	31 (1)
2008	417	31	3 (1)	5 (6)	18	7	6	2 (0)	23 (7)
2009	751	75	3 (1)	6 (6)	11	52	17	3 (1)	36 (16)
2010	1044	129	3 (1)	5 (4)	36	68	28	3 (1)	35 (16)
2011	1158	119	3 (1)	5 (6)	34	61	25	2 (0)	34 (16)
2012	478	37	3 (1)	5 (4)	18	15	6	2 (0)	26 (9)
2013	64	3	3 (0)	3 (1)	1	1	1	2 (0)	17.0 (0)
2014	15011	433	3 (1)	5 (4)	99	292	56	2 (0)	39 (19)
2015	19520	484	3 (1)	5 (4)	115	356	44	3 (1)	50 (36)

## Mapping the diffusion risks of Dengue Fever

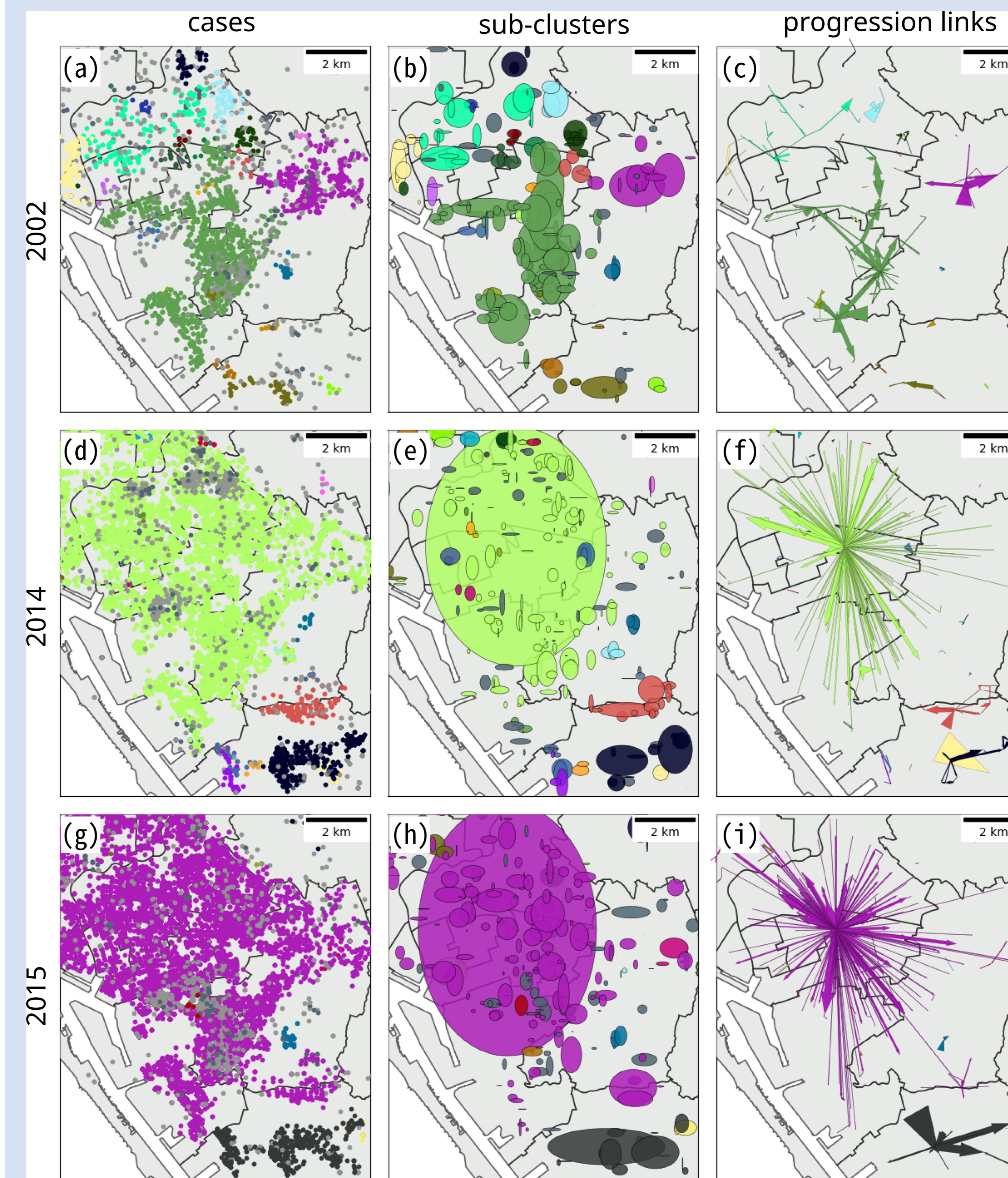


Figure 3: The distribution of cases, sub-clusters, and progression links, in three years.

The spatial distribution of the cases, sub-clusters and progression links in 2002, 2014, and 2015 are shown in Figure 3. The colors indicate the progression chains to which they belong. The sub-clusters are presented as standard ellipses using the number of cases as the standard distance. The width of the progression links indicates the number of shifting links that connect the two sub-clusters. In 2002, the algorithm found a significant spatial separation between the progression chains (Figure 3a). The spatial distribution of the progression chains in 2014 and 2015 were different from the year 2002, in that an extremely large sub-cluster appeared and covered the most highly populated area of the city (Figure 3e and 3h).

The animations of the space-time distributions of diffusion patterns with temporal dynamic is available in the following web-page.  
QR code:



url:  
<http://goo.gl/vZUGw2>

## References

- [1] Tzai-Hung Wen, Chieh-Ting Tsai, and Wei-Chien-Benny Chin. Evaluating the role of disease importation in the spatiotemporal transmission of indigenous dengue outbreak. *Applied Geography*, 76:137–146, nov 2016.
- [2] World Health Organization. Epidemiology, burden of disease and transmission. In *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control*, chapter 1, pages 1–21. World Health Organization, Geneva, 2009.