Using Open Data for Spatial Analysis of Dengue Fever Risk in SINGAPORE

NEW YORK UNIVERSITY

Sarah Welt
M.S. Applied Urban Science and Informatics
Master of Business Administration
Dengue is potentially fatal and global incidents have grown in recent decades making it the leading cause of death among children in some countries.

There is no specific treatment for Dengue, but early detection and access to proper medical care lowers fatality rates significantly.

Assessing risk is essential to making sure prevention and treatment efforts are appropriately targeted.
Using data from Singapore’s open data portal two methodologies were used to quantify risk:

1. spatial buffers measure population at risk
   - Mosquito breeding habitat
   - Dengue Fever Cluster

2. unit weighted regression → risk index
The mosquito breeding habitat data was separated into the five regions:

1. Northeast
2. Northwest
3. Southeast
4. Southwest
5. Central

The data was merged into a single shapefile for use in all parts of the analysis.
Mosquito Breeding Habitat
400 Meter Buffer

• Location of mosquito breeding habitats
• Population data from census by planning subzone.
• 400 meter radius from each breeding habitat was determined and then summed.
• Based on this sum and the total population of Singapore, the percentage of the population living within 400 meters of a mosquito breeding habitat was calculated.

DATA PROCESSING

Buffer Tool: Dissolved ½ mile buffers around mosquito breeding habitats

Spatial Join: Join planning subareas that intersect the buffer and sum

Planning Subareas with Population

Total population in buffer
2,541,510 People live within 400 meters of a mosquito Breeding Habitat

66% of the Population
Singapore Dengue Fever Cluster
400 Meter Buffer

- Location of Dengue Fever clusters
- Population data by planning subzone.
- 400 meter radius from each Dengue Fever cluster was determined and then summed.
- Based on this sum and the total population of Singapore, the percentage of the population living within 400 meters of a Dengue Fever cluster was calculated.

**DATA PROCESSING**

**Buffer Tool:** Dissolved ½ mile buffers around Dengue Fever Cluster

**Spatial Join:** Join planning subareas that intersect the buffer and sum

400 meter buffer Dengue Fever cluster

Planning Subareas with Population

Total population in buffer
888,520 People live within 400 meters of a Dengue Fever Cluster

26% of the Population
\[ \hat{y} = \hat{f}(x) = \hat{b} + \sum_{i} x_i \]

Improper Linear Model
Using information from the World Health Organization four variables were identified as predictive of Dengue Fever:

1. Mosquito breeding habitats
2. Dengue fever clusters
3. Percent of population under 15
4. Population density
METHODOLOGY

1. Calculate Z score to remove scale

2. Convert Features to points
   - Mosquito Breeding Habitats
   - Dengue Fever Clusters
   - Percent of Population under 15
   - Population density
3. Create kernel densities for the 4 inputs with 50sqft grid cells.
METHODOLOGY

4. Added the surfaces weighted by .25 to average them.
**LIMITS OF ANALYSIS**

**400 Meter Buffers:**
- Not every person in the planning subarea may be within the 400 meter buffer.
- If only part of the subzone was within the 400 meter radius, the population for the whole subzone was included.
- The number of people within 400 meters may be overstated.

**Variables:**
- Information from the World Health Organization was taken at face value.
- No attempt was made to verify the predictive nature of the variables used.

**Dengue Fever Risk Index:**
- Only two weeks worth of Dengue Fever cluster data was available for this analysis.
- The methodology of the Dengue Fever Risk Index was chosen to limit the impact of the sparse data.
- Adding data for a longer time period would improve the quality of the index.
- As new cases of Dengue are reported the index can be refined to better predict the real world occurrences.
Thank you!

Full paper available here:
http://tinyurl.com/googledrive-denguefever

iMapOpenData
https://imapopendata.tumblr.com/