

# Spatial Pattern of the Dengue Vector in Iquitos, Peru

## STOR 940 Summary of Final Presentation

Rob Erhardt

### 1 Introduction

Dengue is a disease transmitted by the mosquito *aedes aegypti*, and it affects an estimated 100 million people worldwide each year. About 40% of the world's population live in areas at risk of infection. There is also a more serious form of the infection called Dengue Hemorrhagic Fever (DHF). Even with proper medical care, mortality from DTH is around 1%.

With no vaccine or medicine specifically designed to treat Dengue, attention naturally shifts to prevention. Minimizing the risk of the disease equates to minimizing the exposure to *aedes aegypti*. Getis, Morrison, Gray, and Scott conducted a spatial study of the distribution of the *aedes aegypti* in Iquitos, Peru, a city of about 345,000 people 120m above sea level in the Amazon basin area of Peru. They find strong evidence of clustering for both adult and pupal *aedes aegypti* at a distance of 10m, and weaker evidence that adults cluster with distances as large as 30m.

### 2 Data and Analysis

Data were collected at each house in two distinct neighborhoods of Iquitos, Peru (528 and 481 houses, respectively). At each location, four variables were measured: the number of adult *aedes aegypti*, the number of pupae, the number of contaminated water-holding containers, and the total number of water-holding containers. The survey was repeated 3 weeks later.

To identify clustering, the authors used the point and weighted K-functions on each of the four variables measured. The point K-function estimator is

$$\hat{K}(d) = \frac{A}{N^2} \sum_i \sum_j u_{ij}^{-1} I_d(d_{ij} < d), \quad (1)$$

where  $A$  is the area of study,  $N$  is the number of points,  $d_{ij}$  is the distance between the  $i$ 'th and  $j$ 'th observed point, and  $u_{ij}$  is the border correction (if  $d_{ij} < \text{distance to border of set}$ ,  $u_{ij} = 1$ . Otherwise,  $u_{ij}$  is the proportion of the circumference that lies within  $A$ ). It is convenient to transform the K-function as

$$\hat{L}(d) = \sqrt{\hat{K}(d)/\pi}, \quad (2)$$

because  $\hat{L}(d)$  now has expectation  $d$  and approximate variance  $1/2(\pi N^2)$ .

The generalization of equation (1) to weighted points is the weighted K-function, with transformed estimator

$$\hat{L}_w(d) = \left[ \frac{\left( A \sum_i \sum_j u_{ij}^{-1} I_d(d_{ij} \leq d) x_i x_j \right)}{\pi \left[ \left( \sum_i x_i \right)^2 - \sum_i x_i^2 \right]} \right]^{1/2}. \quad (3)$$

Notice now that terms  $x_i$  and  $x_j$  are present, which correspond in general to the weights of points  $i$  and  $j$ . For the study of *aedes aegypti*, these variables correspond to the number of mosquitoes at each location. Below is an example of the unweighted and weighted L-functions for a random point process on the unit square with counts proportional to  $x$ :

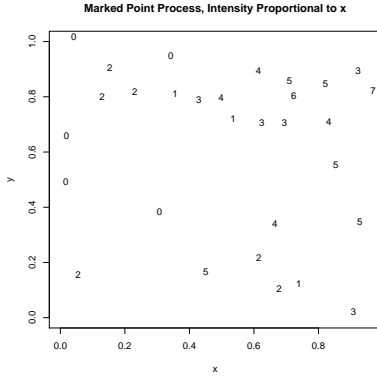


Figure 1: Random locations, counts follow a Poisson process with intensity proportional to  $x$

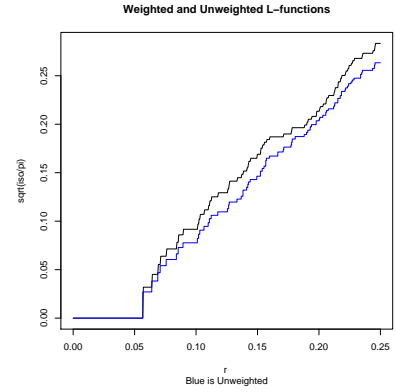


Figure 2: Weighted (Blue) and Unweighted (Black) L-functions.

Once the presence of a particular form of clustering at a particular distance had been established, the next task was identify individual houses as either members or non-members of clusters. This was facilitated using the local  $G_i^*$  statistic, which is

$$\hat{G}_i^*(d) = \frac{[\sum_i w_{ij}(d)x_j - W_i^*\bar{x}]}{s\left(\frac{[NS_{1i}^* - W_i^{*2}]}{(N-1)}\right)^{1/2}}, \quad (4)$$

where  $w_{ij}(d)$  is the matrix of 0/1 entries with  $w_{ij} = 1$  if  $i$  and  $j$  are neighbors at distance  $d$ ,  $W_i^* = \sum w_{ij}(d)$ , and  $S_{1i}^* = \sum w_{ij}^2$ . The local G-statistic is asymptotically normal (Zhang, 2008). For each position  $i$ , the local G-statistic is calculated at each distance  $d$  where significant clustering was found, and the resulting value is compared to a critical value of 2.575 (corresponding to a two-sided hypothesis test at the 0.01 significance level).

### 3 Results

Table 1 shows the clustering for houses and for adult mosquitoes

Distance (meters)	House	Adult mosquitoes	House Increment	Adult Increment
10	16.33	22.86	13.33	22.86
20	27.13	36.79	10.80	13.93
30	38.70	50.58	11.57	13.79
40	52.85	61.13	14.15	10.55
50	65.67	74.24	12.82	13.11
60	76.70	83.94	11.03	9.70
70	88.03	93.71	11.33	9.77
80	100.98	104.12	12.95	10.41
90	111.77	113.10	10.79	8.98
100	122.19	120.57	10.42	7.47

It is easiest to look at the increment columns to interpret the results. Since we expect the L-function to equal  $d$ , we expect each entry in the increment columns to equal 10. Entries significantly above 10 indicate clustering, while those below 10 indicate dispersion. At  $d = 10$ , there is mild evidence that houses are clustered (13.33 vs. 10), which is not surprising. However, there is even stronger evidence that adult mosquitoes are clustered (22.86 vs. 10), well above the clustering due to houses. To attach statistical significance to this result, the authors took the positions of the houses as fixed, and permuted the numbers of mosquitoes under the null hypothesis of complete spatial randomness. This procedure was repeated 19 times, and for each permutation the value of  $\hat{L}(d = 10)$  was calculated, resulting in bounds of (11.88, 19.10). The observed value of 22.86 falls well outside of this interval,

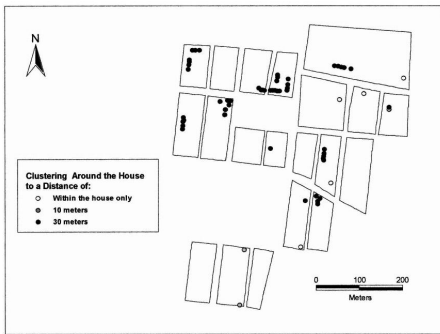


FIGURE 3. Clusters of *Aedes aegypti* adults in the Maynas  $\alpha$  study based on the number of mosquitoes in houses.

which the authors take to indicate a significant result at the  $\alpha = 0.05$  level.

Weaker evidence of clustering of adult mosquitoes exists at distances up to 30m (adult increments are 13.93 and 13.79 respectively). Clearly, had the authors repeated the permutation tests for  $d = 20$  and  $d = 30$ , they would not have found statistically significant results. Still, they claim that this result shows weak evidence for clustering. The only other significant result of clustering was pupae at a distance of 10m. No significant evidence of clustering was found for larger distances, or for the remaining variables (contaminated water-holding containers and all water holding containers).

Having established several types of clustering, the local G-statistic was used to identify individual members of clusters. First, for adult mosquitoes the authors calculated  $G_i^*(d)$  for  $d=1m, 10m, 20m,$  and  $30m$ . G-statistics in excess of 2.575 were taken as significant at the 0.01 level. They repeated this calculation for pupae, again taking the same distances and significance level. Results are best shown in an image. The figure above shows clustering of adult mosquitoes in the Maynas neighborhood.

## 4 Discussion

This study supports using the household as the basic unit for entomological surveys. It superficially supports focal insecticide treatment where all households within 50-100 m are treated with insecticide following a confirmed case of dengue, with three shortcomings:

1. One must study time stability of the spatial clustering, since there is a delay between being bitten and being diagnosed
2. Infected people can themselves transmit the disease over far greater distances
3. People may not seek medical attention

Secondary goals of studying the time-stability of the spatial pattern, as well as the correlation between pupal and adult clustering, were not fully achieved in this study, and the authors mention future work that will better address those research areas.

## 5 References

1. Getis, Morrison, Grey, Scott, 2003. Characteristics of the Spatial Pattern of the Dengue Vector, *Aedes Aegypti*, in Iquitos, Peru. *American Journal of Tropical Medicine and Hygiene*
2. Zhang Tonglin, 2008. Limiting Distribution of the G Statistics. *Statistics and Probability Letters*