A geographical sampling strategy for field surveys in an urban area using high-resolution satellite imagery

Adriana Troyo[1,2], Douglas O. Fuller[3], Olger Calderón-Arguedas[2], and John C. Beier[1]

[1]Global Public Health Program, Department of Epidemiology and Public Health, University of Miami. [2]Centro de Investigación en Enfermedades Tropicales, Departamento de Parasitología, Facultad de Microbiología, Universidad de Costa Rica. [3]Department of Geography and Regional Studies, University of Miami

Abstract

Field evaluations for studying the epidemiology of vector-borne diseases like dengue in urban areas are often restricted to selection of households and buildings for field surveys. Therefore, the resulting sampling frame may exclude specific locations within the urban environment that contain vector habitats and thus may bias the results. A sampling strategy was developed for field surveys in an urban area using high-resolution satellite imagery. The site selected was Puntarenas, a city affected by dengue on the Pacific coast of Costa Rica, for which high-resolution satellite imagery was available from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER, 15 m spatial resolution) and QuickBird (0.6 m and 2.4 m spatial resolution for panchromatic and multispectral bands, respectively). Grids obtained from the ASTER imagery and a cover map generated from the QuickBird multispectral bands were used to determine the optimal grid area of 10 000 m^2, which contain 13±6 houses. A final grid 42 by 42 pixels (100.8 x 100.8 m) was created using the multispectral Quickbird imagery, and cells that had an area less than 90% within one specific locality of Puntarenas were excluded. The remaining cells were grouped according to locality and a random sample (10%) was selected from each. This sample of cells would be used for field data collection on specific mosquito larval habitats by evaluating the entire area within the geographical limits of each cell. To assess the suitability of the selected grid cells, the proportion of tree area ("tree" class Kappa = 0.91) was extracted for the individual cells from the QuickBird cover map. The mean percentage of tree cover in each locality and total area was compared between the selected sample cells and the total cells of the Puntarenas image. Overall, the sample adequately represented the total area and most of the individual localities in terms of tree cover. In 8 of 10 localities the difference between the estimate (sample) and the real percentage of tree cover was less than 3%. These results show that high-resolution satellite imagery and geographical information systems are useful in evaluating urban areas and randomly selecting sections for field data collection on mosquito larval habitats that are practical, representative, and will reduce bias.

Field evaluations for studying the epidemiology of diseases in urban areas are often restricted to sampling of households and buildings during surveys. The resulting sampling frame may exclude locations within this complex environment that would provide valuable information and thus bias the results. Therefore, different sampling strategies may need to include non-residential sites in entomological field surveys, such as those required for studying dengue[1] and other vector-borne diseases of urban environments.

Remote sensing and geographical information systems (GIS) have been used to study the epidemiology of vector-borne diseases.[2,3] These technologies offer powerful tools for describing, explaining, and predicting epidemiological phenomena, which can be used to develop or improve surveillance, prevention, and control strategies.[4] High resolution data currently available is useful for studying factors that affect diseases within the urban environments. In this study, sampling strategy was developed for the Great Puntarenas area, Costa Rica, using high-resolution satellite imagery and GIS technology.

Materials and Methods

The site selected to develop the sampling method was Puntarenas, a city affected by dengue on the Pacific coast of Costa Rica. High-resolution satellite imagery from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER, 15 m spatial resolution) and QuickBird (0.6 m and 2.4 m spatial resolution for panchromatic and multispectral bands, respectively) was available for this area. A land cover map was generated from the QuickBird imagery by using the automated back propagation neural network in Idrisi Kilimanjaro software. Grids of different sizes were obtained from the ASTER imagery to estimate the number of houses/small buildings per area of the land cover map and to determine an optimal size for the cells to be sampled. A final grid was created using the multispectral Quickbird imagery, and cells were grouped according to locality. A random sample of cells was selected from each locality, which was proportional to the total number of cells, to ensure at least one representative sample set of each locality. To initially assess the suitability of the selected sample grid cells, the QuickBird land cover map was used to extract the proportion of tree area ("tree" class Kappa = 0.91) in individual grid cells, as well as in the total area of the localities. For each locality, the mean percentage of tree cover in the selected sample cells was compared to the mean percentage of tree cover in the total cells and the percentage of tree cover in the total area of the locality.

Results

This sampling method was intended for field data collection on specific mosquito larval habitats. According to the mean number of houses per area, an optimal grid cell area that would be operationally adequate was estimated at 10 000 m^2. This cell size of 100 by 100 m contained 13±6 houses (Fig. 1), and was considered large enough for a team of 2 people to search in half a day (approximately 3 hours at 15 minutes per house). The final grid created from the multispectral Quickbird imagery contained cells 42 by 42 pixels (100.8 x 100.8 m), and only the cells that had more than 90% of their area within one specific locality of Puntarenas were included in the sampling frame (Fig. 2). The random sample consisted of 36 cells, approximately 10% of the total 355 cells (Fig. 3). This number of grid cells selected was such that the time taken to collect the field data would not exceed 3 weeks, since it is necessary for the data to be analyzed within approximately homogeneous external environmental conditions of each season. In 8 of 10 localities the difference between the estimated percentage of tree cover (from sample calls) and the real percentage of tree cover was less than 3%. By displaying the cells on the QuickBird panchromatic image, small features that serve as visual limits for field observations like roads, houses, and trees can be identified (Fig. 4).

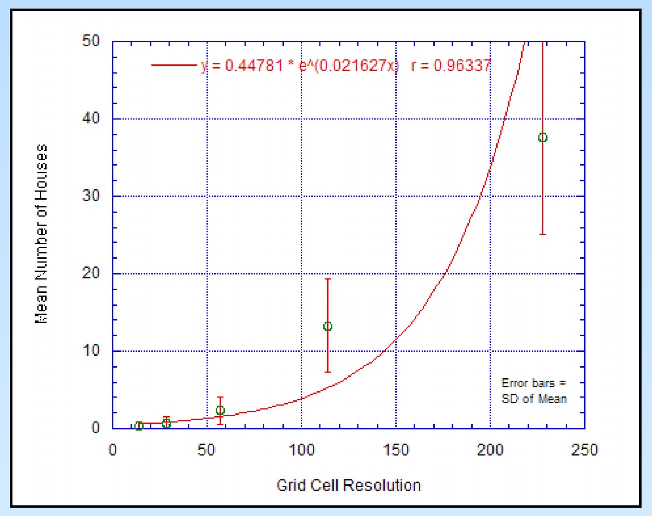

Fig. 1. Sample grid cell size and mean house numbers for Puntarenas.

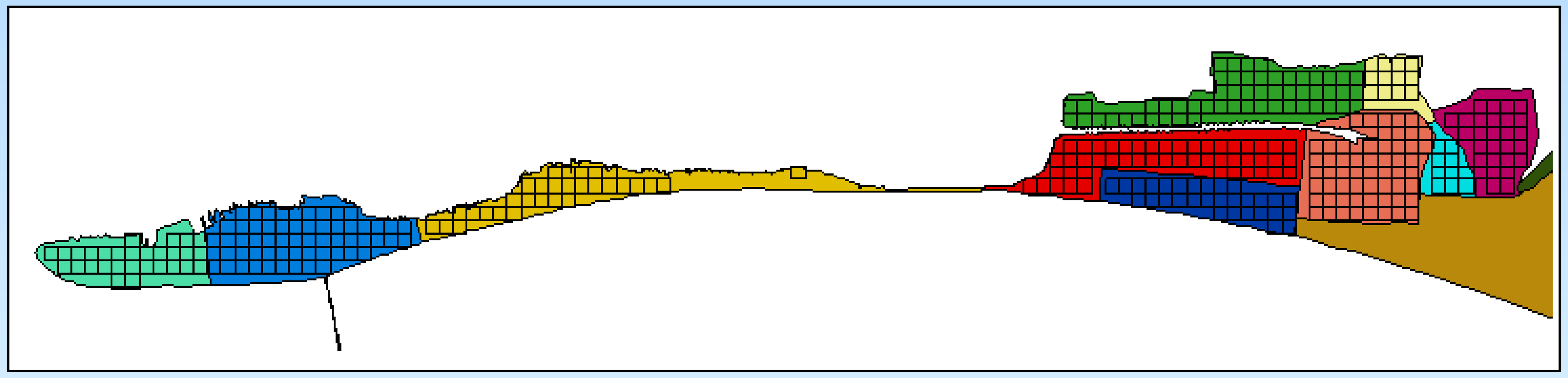

Fig. 3. Sampling frame composed of grid cells within the individual localities of Puntarenas.

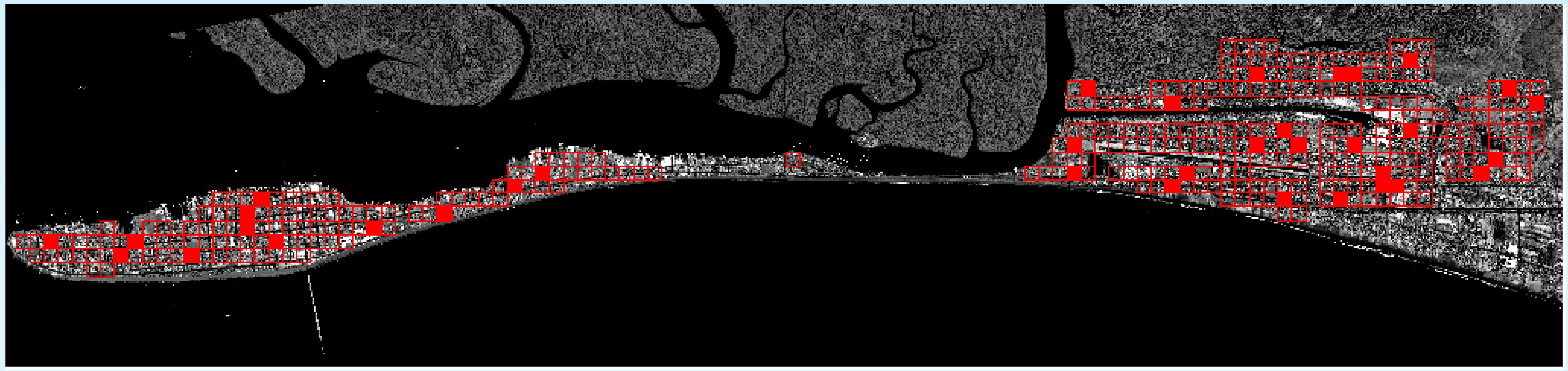

Fig. 2. Random sample of grid cells (10%) selected for field studies in Puntarenas.

Table 1. Comparison of the average tree cover values extracted from the randomly selected cells and those of the total cells that represent the area.

Locality	% tree of total area	% tree mean of total cells	% tree mean of sample	Standard Error	Difference
Barrio El Carmen	22.62	25.15	22.15	5.51	-3.00
El Centro	15.36	16.35	17.75	3.84	1.40
El Cocal	18.82	19.07	29.74	5.80	10.67
Veinte de Noviembre	41.46	40.60	49.31	8.05	8.71
Chacarita	26.51	20.82	18.84	5.96	-1.98
Fray Casiano	41.82	38.28	38.55	4.77	0.27
San Luis	50.53	48.09	47.85	9.54	-0.24
Carrizal	40.37	41.38	42.48	9.58	1.10
El Huerto	54.82	54.38	56.75	19.61	2.37
Linda Vista	54.36	50.21	48.47	12.46	-1.74
Total area	33.60	32.97	35.08	3.20	2.11

Fig. 4. Sampled cell of Puntarenas showing detailed structures of the area to be analyzed.

Conclusions

Remote sensing and GIS technology provided useful methods to develop a sampling frame for field studies within urban Puntarenas. Although coarser resolution satellite imagery have been used to develop similar entomological sampling methods for malaria vectors,[5] the method presented here shows that detail provided by high-resolution satellite imagery allows more precise calculations of optimal cell size, as well as useful information for planning operations previous to the site visit. Although high-resolution satellite imagery and GIS were used to evaluate urban areas and randomly select sections aimed at obtaining data on mosquito larval habitats, this method can be applied to sample other interactions and disease systems in urban environments. These strategies would reduce biases and provide information from the field that is both practical to obtain and representative.

Selected references:

1. Morrison AC, Sihuincha M, Stancil JD, Zamora E, Astete H, Olson JG, Vidal-Ore C, Scott TW. 2006. *Aedes aegypti* (Diptera:Culicidae) production from non-residential sites in the Amazonian city of Iquitos, Peru. *Ann Trop Med Parasitol Suppl 1:*73-86.
2. Hay SI, Packer MJ, Rogers DJ. 1997. The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *Int J Remote Sensing 18:*2899-930.
3. Beck LR, Lobitz BM, Wood BL. 2000. Remote Sensing and Human Health: New Sensors and New Opportunities. *Emerg Infect Dis 6:*217-26.
4. Rogers DJ, Randolph SE. 2003. Studying the gobal distribution of infectious diseases using GIS and RS. *Nature Rev Microbiol 1:*231-7.
5. Keating J, Macintyre K, Mbogo C, Githeko A, Regens JL, Swalm C, Ndenga B, Steinberg LJ, Kibe L, Githure JI, Beier JC. 2003. A geographic sampling strategy for studying relationships between human activity and malaria vectors in urban Africa. *Am J Trop Med Hyg 68:*357-65.

This research was supported by NIH grant 1 P20 RR020770.