Research article

Fuzzification of multi-criteria proxy geoclassifiable vegetation and landscape biosignature estimators to predict the potential invasion of *Aedes aegypti* in Barcelona, Spain

Patricia Leandro-Reguillo\textsuperscript{a}, Toni Panaou\textsuperscript{b}, Ryan Carney\textsuperscript{c,d}, Benjamin G. Jacob\textsuperscript{b}\textsuperscript{*}

\textsuperscript{a}Department of Environmental Health, College of Public Health. University of South Florida, Tampa, FL, USA.
\textsuperscript{b}Department of Global Health, College of Public Health. University of South Florida, Tampa, FL, USA.
\textsuperscript{c}Department of Integrative Biology, College of Arts and Sciences. University of South Florida, Tampa, FL, USA.
\textsuperscript{d}Center for Virtualization & Applied Spatial Technologies, College of Arts and Sciences. University of South Florida, Tampa, FL, USA.

E-mail: bjacobi@health.usf.edu, apanaou@mail.usf.edu, patricia.sensibilidad@gmail.com

*Corresponding author. E-mail:bjacobi@health.usf.edu

This work is licensed under a Creative Commons Attribution 4.0 International License.

Abstract

After decades of absence, *Aedes aegypti* mosquito has returned to European domains. This urban species - major vector of Dengue, Chikungunya, and Zika virus - may pose a major public health hazard if authorities do not halt its expansion over this continent. Therefore, it is necessary to target potential breeding areas within Mediterranean temperate cities to improve the effectiveness of entomological surveillance and vector source reduction programs. Due to its intense commercial and passengers flux, the port city of Barcelona was considered most at risk of *Ae. aegypti* re-introduction. In order to assess the areas of the city with the highest possibility of breeding site concentration, a multi-criteria spatial model was constructed. A fuzzy overlay analysis was implemented for four, vector, urban-reproduction, risk factor geolocations: narrow streets, high population density, urban vegetation and courtyards. Results highlight three main risk areas, Ciutat Vella (historic centre), Sants and El Clot neighborhoods; being the historic centre most at risk due its seaport proximity. Public health decision-making and interventions may be enhanced by considering the use of multi-criteria spatial modeling...
freely available satellite images and fuzzy logic software for predictive habitat mapping \textit{Ae. aegypti} spatial distribution patterns.

\textbf{Keywords}: \textit{Aedes aegypti}; Barcelona; multi-criteria spatial modeling; fuzzy overlay analysis; urban breeding sites.

\section{Introduction}

The expansion of \textit{Aedes egypti} mosquitoes in Europe could pose a major public health threat on this continent. \textit{Ae. aegypti} is considered a major vector of diseases such as Yellow Fever, Dengue, Chikungunya and Zika (Akiner \textit{et al}. 2016). Current European vector control interventions and surveillance programs could be significantly improved by implementing targeted geomorphic, predictive, \textit{Ae. aegypti} habitat epidemiological, risk models in decision making.

During the late nineteenth century until the 1950s, Yellow Fever and Dengue outbreaks were common along the temperate Mediterranean coast, due mainly to the introduction and spread of this species from Africa through sailing ships (Schaffner and Mathis 2014). However, post World War II, the vector mosquito disappeared from this geographic area; an event that is not clearly explained (Akiner \textit{et al}. 2016).

Currently, Europe is closely monitoring two well established populations of \textit{Ae. aegypti} that appeared in the Atlantic isle of Madeira and Turkey in 2004 and 2008 respectively (Akiner \textit{et al}. 2016), areas that were previously geographically unsuitable for the vector. Due to the intense commercial and passenger flux that Europe maintains both internally and with the Americas, the World health Organization (WHO) and the European Centre for Disease Prevention and Control (ECDC) have recommended the implementation of vector control-source reduction programs and entomological surveillance throughout the continent to hamper the vector’s expansion (ECDC 2016, WHO 2016). These actions were successfully applied in The Netherlands in 2010, after the discovery of \textit{Ae. aegypti} mosquitoes that were introduced in the country through a shipment of tires from Miami, USA (Brown \textit{et al}. 2011).

However, these programs were not generally implemented in all European countries. According to the ECDC’s known distribution map of \textit{Ae. aegypti} mosquitoes in Europe in April 2017, few countries possess data throughout all its territory, including Spain (ECDC 2017). This country is one of the European areas with highest risk of imported cases of Zika due its elevated commercial and passenger flux from South and Central America (Rocklöv \textit{et al}. 2016). According to the Spanish Ministry of Health, from 2015 to May 2017 there were
321 recorded cases of Zika virus, all imported except two cases infected within the territory via sexual transmission (Ministerio de Sanidad, Servicios Sociales e Igualdad 2017). Since *Aedes albopictus* mosquitoes - present in this country - have not yet shown effective for Zika transmission, the re-introduction of *Ae. aegypti* could pose a major public health threat.

Under this scenario, prediction spatial models may analyze topographic and urban patterns for potential breeding sites of *Ae. aegypti* which may be useful in decision making to detain its propagation. The use of geographic information systems (GIS) has risen as one of the most powerful tools in vector control. High-resolution satellite imagery and Lidar (light detection and ranging) data provide researchers with high quality remote sensing information that combined with GIS software allow the creation of improved predictive spatial models in rural and urban environments (Delmelle *et al.* 2014, Fuller *et al.* 2010, Kolivras 2006, Troyo *et al.* 2008, Yoo 2014, Zou *et al.* 2016).

In Jacob and Novak (2016) a predictive spatial C++ model was constructed to epidemiologically forecast, Yellow Fever case distribution employing geographically and entomologically sampled, *Ae aegypti*, breeding foci, time series information for an agro-irrigated, riceland, village eco-epidemiological, ecosystem, study site in Gulu, Uganda. Initially, a grid-stratified, panchromatic, QuickBird 0.61m, spatial resolution, uncoalesced dataset of endmember, transmittance, land use land cover, (LULC), spectral signatures of the study site was differentitated by visible and near-infrared, (NIR) frequency, wavelength, irradiance values. The object-based classification algorithm employed a divergence measure to match mixed pixels (“mixels”) to a dataset of derivative, end member, LULC spectra. A mosaiciked dataset was then employed to manage, display, serve, and distribute, raster, time series, discontinuous data of intermittent, forest-canopied, *Aedes aegypti*, oviposition, capture point, larval habitat georeferenced, geolocations along the peripheral, riceland, forest-canopied, interface corridor. A new mosaiciked dataset was created as an empty container in the geodatabase cyber-environment with default properties which was subsequently added to the raster dataset. Thereafter, the virus data was cartographically illustratable as non-mixelated, riceland, agro-village, non-homogeneous, gridded, LULC eco-zones (e.g., sparsely shaded, pre-flooded, densely canopied, post-tillering) within a supervised classification matrix in ArcGIS. Interfaces of the geoclassified landscape were noted. Non-normalized, geometric, seasonal, case distribution data was articulated from various spatial perspectives (e.g., iterative Bayesian). The models took inputs from table file in the geodatabase which subsequently imput information of yearly population and yellow fever cases from 1990 to 2012. Log-likelihood functions were also generated in PROC REG for each geosampled, diagnostic, time series, *Ae.aegypti* endmember, sub-mixel, geoclassified,
LULC-specified, visible and NIR, explantorial regressor while the joint log-likelihood function was exponentially logarithmically quantitated.

Spatial autoregressive moving average models were then extended to autoregressive, integrated, moving average, models in PROC ARIMA for conducting, observation-driven, sub-pixel, endmember modelling of any non-Gaussian, non-stationary, time series, *Ae aegypti* geosampled, count variables. A negative binomial, mean model fit was compared to that of a Bayesian Gaussian fit employing Box-Cox transformed data. For the Gaussian the posterior distributions appeared to be platykurtic. The logarithm of the posterior density was calculated. Geospatial outliers were teased out in the residual plots. The normalized modes were validated employing a Monte Carlo simulation. Subsequently, the *Ae aegypti* data was ecologically forecasted in C++ employing a stochastic-dynamic, random, weighted matrix. The model data included, LULC and meteorologically-derived, probabilistic, discrete, integer values which then were employed to prognosticate population growth and number of case occurrences to 2020 in C++ for the Gulu study site.

*Ae. aegypti* mosquito is highly adapted to urban milieus; it is able to breed in small containers, vases, flower pot dishes, tree holes, rock pools, street and patio drains, etc. (Schaffner and Mathis 2014). Therefore, potential breeding sites in Mediterranean cities will be characterized by three main variables: narrow streets, high dense populated neighborhoods, urban vegetation and the presence of courtyards or patios within building blocks. Barcelona is a large coastal metropolitan area that possesses one of the highest commercial and travelers flux of the Mediterranean Sea, is characterized by the Cerda’s courtyard blocks typology (Figure 1), which is featured in large parts of the city (Urbano 2016).

![Figure 1](image-url): Cerda’s courtyard block typology with urban vegetation, narrow streets and high dense populated neighborhoods in Barcelona, Spain. Source: Google Earth 2017.
This study develops a multi-criteria spatial model that displays areas with the highest breeding potential for *Ae. aegypti* in Barcelona. The model will help us to predict future spatial-temporal distribution of this species throughout the city especially if we consider its re-introduction via maritime traffic (Figure 2) during suitable months for vector reproduction (June, July and August) (Rocklöv et al. 2016).

![Port of Barcelona](image)

**Figure 2:** Port of Barcelona. Source: Google Earth 2017.

**B. Materials and methods**

**B.1. Resources**

Georeferenced datasets employed in this research were: sub-parcel cartographic maps for four districts of the Barcelona: Ciutat Vella, L’Example, Sants-Montjuic and Sant Martí (Figure 3); released by the City Plan Department of the Barcelona City Council in 2017.
Figure 3: Districts of interest within the street network map of Barcelona: (1) Ciutat Vella, (2) Example, (3) Sants-Montjuic and (4) Sant Marti.

Multispectral band Sentinel-2 satellite imagery with 10 m of size resolution and dated on 2017-06-22 was obtained from the European Spatial Agency (ESA 2017). Lidar data was also obtained from the Geological and Cartographic Institute of Catalonia (2017). Lidar points were distributed in LAS 1.2 files containing blocks of 2 x 2 km and with a minimum density of 1.5 points/m². The blocks obtained for this work are shown in Figure 4.
B.2. Density and NDVI raster maps

For this study, a multi-criteria spatial modelling and the fuzzy overlay function was adopted to create a distribution map of potential, urban, breeding sites, areas for *Ae. aegypti* in four districts of Barcelona. Although the overlay analysis method has been used in decision making mainly for site selection and suitability projects (Esri 2016, Baidya et al. 2014), the objective of this work was to employ this methodology as a useful tool in vector control-source reduction programs and entomological surveillance. The general process followed for the overlay analysis is graphically explained in Figure 5.

**Figure 4**: Gridded map depicting Lidar data blocks used in this study.

**Figure 5**: Flow chart for fuzzy overlay analysis
To perform an overlay analysis it is necessary to measure a variety of inputs (www.esri.com). Following the breeding pattern of this mosquito, four urban LULC characteristics of populated metropolitan areas were identified as potential breeding sites: drains in narrow streets, courtyards within building blocks, population density and vegetation. Accordingly, four raster maps were developed using ArcMap 10.5.

First, three raster density maps were created based on cartographic information obtained from the Barcelona City Council: street network density, courtyards density and building density (see Figure 6). Line density tool, used to generate the road density map, yielded the number of polylines per unit area that contained a circumference with radius $= \text{x}$ shaped around the cell. Point density tool, used to create courtyards density and building density, geoclassified, LULC maps, featured the number of points (obtained from polygons) included within a circumference with radius $= \text{x}$ shaped around the cell.

**Figure 6:** Density maps process and fuzzification.
Finally, a normalized difference vegetation index (NDVI) map of Barcelona was generated through near-infrared NIR light and visible (Red) spectral bands included within the composite raster image obtained from Sentinel-2 multispectral bands. The NDVI is a simple graphical indicator that can be employed to analyze remote sensing measurements, typically but not necessarily from a space platform, and assess whether the target being observed contains live green vegetation or not (www.esri.com). The appeal of NDVI is its simplicity and its relationship either empirically or theoretically to biophysical LULC variables (Bannari et al. 1995). NDVI's have been proven to be well correlated with various vegetation parameters such as green biomass (Tucker et al. 1986), chlorophyll concentration (Buschmann and Nagel 1993), leaf area index (LAI) (Asrar et al. 1984), foliar loss and damage (Olgemann 1990), photosynthetic activity (Sellers 1985) and carbon fluxes (Tucker et al. 1986).

The NDVI is a widely used vegetation index in arboviral mosquito epidemiology (Jacob et al. 2009, Brown et al. 2008, Cooke et al. 2006, Kunkel and Novak 2005, Ward et al. 2004, Backenson et al. 2002, Brownstein et al. 2002, Linthicum et al. 1987). For example, Brown et al. (2008) employed canonical correlation analyses to determine if a significant relationship existed between NDVI, disease/water stress index and distance to water and four local West Nile Virus (WNV) competent vectors (Culex pipiens, Culex restuans, Culex salinarius, and Ae. vexans). Their model determined a significant relationship existed between the sampled explanatory predictor covariates and the sampled mosquito habitats (0.93, P=0.03). The final model outputs explained 86% of the variance in the environmental and mosquito measures. Diuk-Wasser et al. (2006) constructed regression models to predict high and low adult mosquito abundance sites for determining arboviral activity in Fairfield County, Connecticut USA. The best predictive models included non-forested areas for Cx. pipiens, surface water and distance to estuaries for Cx. salinarius, surface water and grasslands/agriculture for Ae. vexans, distance to palustrine habitats for Culiseta melanura, and, seasonal difference in the NDVI parameters. The best models included non-forested, grid-stratified LULC areas for Cx. pipiens, surface water and distance to estuaries for Cx. salinarius, surface water and grasslands/agriculture for optimally quantitating Ae. vexans and seasonal difference in the NDVI and distance to palustrine habitats for Cs. melanura.

The equation used to generate the NDVI map of the Barcelona study regions was:

\[
NDVI = \frac{P_{\text{red}} - P_{\text{NIR}}}{P_{\text{NIR}} + P_{\text{red}}}
\]

Results obtained from this formula were employed to provide a discrete integer value per raster cell that ranges from -1 to +1. No deduced, vegetation, grid-stratified, LULC raster cell rendered values close to zero, meanwhile vegetation areas values will approach to +1.
B.3. Fuzzy overlay methodology

Since the four raster map possessed different scale values, it was necessary to conduct a transformation and homogenization of these values in order to develop the final multi-criteria map. Therefore, the fuzzy logic for overlay analysis was implemented through the fuzzy membership process and fuzzy overlay analysis. Fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number which may be employed to handle the concept of partial truth, where the truth value may range between completely true and completely false (Novák et al. 1999).

The fuzzy logic defined the possibility (not probability) that a value is in the raster dataset employing a finite set $X$ where the fuzzy set equation $A$ on $X$ was expressed as following:

$$A = \frac{\mu_1(x_1)}{x_1} + \frac{\mu_2(x_2)}{x_2} + ... + \frac{\mu_n(x_n)}{x_n} = \frac{\mu(x)}{x_1},$$

where $\mu(x)/x_i$ was the membership value to fuzzy set $A$ for $x_i$ (see $\sum_{i=1}^{n}$ Kainz W. n.d.).

In every raster map, transformed values of each raster cell, was defined by the ArcMap linear Fuzzy Membership function, (i.e., fuzzification process) which ranged from 0 to 1. The value 1 indicated certainty that the cell was in the raster dataset and 0 indicated certainty that the cell was not in the set (Figure 7).

![Figure 7: Variations of the fuzzy linear membership function.](Image)

Once all four set explanatory values were homogenized, the Fuzzy Overlay function was run. The Fuzzy Overlay not only analyzed the possibility that certain data feature attributes appeared in multiple georeferenceable inputs; the overlay analysis also assessed the relationships existing among the memberships of multiple variable inputs. To determine the areas with the maximum values of membership in all four criteria sets the FuzzyOr overlay function was performed. This function followed the equation:

$$\text{fuzzyOrValue} = \max(\arg 1, \ldots, \arg n)$$
Thereafter, Lidar cloud points were employed to create a digital elevation model (DEM) slope coefficients and contour lines from ground class 2 data in order to build a the final 3D model in ArcScene 10.5. Automated generation of drainage networks has become increasingly popular with the use of GIS and availability of digital elevation models (DEMs) (www.esri.com). These models account for topographic variability and their control over soil moisture heterogeneity and runoff within a shed. Jacob et al. (2010) evaluated environmental factors such as elevation range to determine human Eastern Equine Encephalitis Virus (EEEV) risk in Tuskegee, Alabama. The model yielded several catchment eco-hydrological explanators including percent surface saturation and total surface runoff for identification of potential productive Culex erraticus aquatic larval habitats, a mosquito vector of EEEV. Results of the DEM analyses indicated a statistically significant inverse linear relationship between total sampled mosquito data and elevation ($R^2 = - .4262; p < .0001$), with a standard deviation (SD) of 10.46, and total sampled bird data and elevation ($R^2 = - .5111; p < .0001$), with a SD of 22.97. DEM statistics also indicated a significant inverse linear relationship between total sampled Cx. erraticus data and elevation ($R^2 = - .4711; p < .0001$), with a SD of 11.16. Here we combined the resulting Fuzzy Overlay DEM with the topography of the Barcelona study site.

### 3. Results and Conclusions

In this work fuzzified maps detailed potential Ae. aegypti breeding sites areas for each predictor variable tested within neighbourhoods geolocated nearby seaports in Barcelona. Since the flight range of these species is limited, an average of 400m according to WHO (2017), density value was considered a suitable parameter in forecasting mosquito habitat distribution. The building density map located zones that possessed an elevated geoclassifiable, LULC built area. Using these variables as an indicator of population density in mixed use districts in Barcelona, the fuzzified building density map revealed that the historic centre and west area have a possible high population density of potential habitats (value=1) (Figure 8).

The courtyard or patio density map revealed the distribution and concentration of this Mediterranean feature along our area of study. Since it was chosen as an indicator of location of small containers, vases, flower pot dishes and patio drains, the fuzzified patio density LULC map located the possible accumulation of these potential breeding sites. We observed that the possible highest density of potential habitat occurs mainly in the first mayor development of Barcelona (Figure 9), characterized by Cerda’s courtyard blocks typology (see Urbano, 2016).
Our third variable, road density, was selected as a predictive indicator of narrow streets location in Barcelona. Road drains geolocated along these usually shaded and low motor traffic streets are suitable breeding habitats for *Ae. aegypti* mosquitoes. The fuzzified road density map revealed the possibility of highest concentration of narrow roads mainly within the historic centre (Figure 10).

The last urban explanatory variable for mosquito reproduction assessed in this study was urban LULC vegetation. This included not only urban parks but also road vegetation such as bushes and trees. The NDVI map geolocated these areas and the fuzzified NDVI map revealed the possibility of urban geoclassified LULC vegetation in Barcelona. The mayor concentrations of urban green areas are Montjuic – just in front the commercial port- and Ciutadella Park – nearby the historic centre (see Figure 11).

![Figure 8: Building density and building density fuzzification maps in selected districts of Barcelona.](image)
Figure 9: Courtyard-patio density and courtyard-patio density fuzzification maps in selected districts of Barcelona.

Figure 10: Road density and road density fuzzification maps in Barcelona.
In order to create a final map to help in entomological surveillance decision making, all fuzzy maps developed were combined to create a multi-criteria spatial model of the geoclassified LULC areas with the highest possibilities of the occurrence of potential *Ae. aegypti* breeding sites within the study site. In this map four sea ports of Barcelona were geolocated, which were deduced as probable points of entry of the vector into the city: two recreational ports northward and the commercial and touristic ports southward (Figure 5).

Our final map shows three main areas with the highest possibilities of vector reproduction: the historic centre (Poble Sec, Raval and Barri Gotic), Sants neighbourhood (east) and El Clot neighbourhood (north) (Figure 12). Based on the topography of this city -with Montjuic mount limiting access from southern seaports (Figure 13), and the geolocation of the historic centre nearby a sea port, we considered this district with the highest risk of re-introduction and spread of *Ae. aegypti* in Barcelona.

**Figure 11**: NDVI and fuzzified NDVI maps in Barcelona.
Figure 12: Potential *Ae. aegypti* breeding sites (1. historic centre, 2. Sants, 3. El Clot) and seaports as probable vector re-introduction.

Figure 13: Port of Barcelona 3D map: potential *Ae. aegypti* breeding sites and elevation.
In this work we employed a multi-criteria spatial modelling in assessing a future possible scenario (e.g., geolocations of breeding hotspots): the risk of distribution and infection after a possible re-introduction of *Ae. aegypti* mosquitoes into a georeferenceable geolocation in Barcelona; an area where currently imported cases of Zika virus occur but no suitable vector for its transmission. We calculated fuzzy scores for these layers employing ArcGIS and produced overlays where the analysis determined the risk due to the breeding potential of *Ae. aegypti* based on numerous datasets, (i.e. a multi-criteria overlay investigation). The fuzzy overlay combined all possible layers (e.g., NDVI), influencing the *Ae aegypti* breeding foci by analyzing the spatial associations between urban habitat parameters. We were able to determine if *Ae. aegypti* landed via maritime traffic that they could possibly first breed and spread within the historic centre, and possibly create a risk of Zika virus transmission in this area. Consequently, we recommend the implementation of entomological surveillance and vector source reduction programs in this district.

Public health interventions for arboviral infection processes can be enhanced by remotely targeting efforts and resources in specific geoclassified LULC risk areas. We recommend the use of GIS multi-criteria spatial modelling, and the fuzzy overlay function, for the prediction of *Ae. aegypti* spatial distribution. Developing countries, where resources are limited, will particularly benefit from GIS methodologies not only for cartographically representing rural LULC areas but also in highly dense, urban, geoclassifiable, LULC areas and growing cities, where ento-ecological, bio-geographical and urban factors overlap in many different ways. The spatial, epidemiological, GIS, forecast modelling of *Ae. aegypti* distribution should be based on a combination of LULC variables that may influence the prevalence and incidence of arboviral diseases. In doing so public health professionals may better understand potential intervention sites which can aid in designing more effective vector control interventions and surveillance programs.

**References**


http://dx.doi.org/10.1080/01431160802549195

Kainz W., n.d. *Fuzzy logic and GIS* [online]. University of Vienna, Austria. Available from:
http://homepage.univie.ac.at/wolfgang.kainz/Lehrveranstaltungen
/ESRI_Fuzzy_Logic/File_2_Kainz_Text.pdf [Accessed 17 June 2017].


http://www.who.int/denguecontrol/mosquito/en/